



RAT-RS: a reporting standard for improving the documentation of data use in agent-based modelling

Sebastian Achter, Melania Borit, Edmund Chattoe-Brown & Peer-Olaf Siebers

To cite this article: Sebastian Achter, Melania Borit, Edmund Chattoe-Brown & Peer-Olaf Siebers (2022) RAT-RS: a reporting standard for improving the documentation of data use in agent-based modelling, International Journal of Social Research Methodology, 25:4, 517-540, DOI: [10.1080/13645579.2022.2049511](https://doi.org/10.1080/13645579.2022.2049511)

To link to this article: <https://doi.org/10.1080/13645579.2022.2049511>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 30 Mar 2022.



Submit your article to this journal [↗](#)



Article views: 436



View related articles [↗](#)



View Crossmark data [↗](#)

RAT-RS: a reporting standard for improving the documentation of data use in agent-based modelling

Sebastian Achter ^a, Melania Borit ^b, Edmund Chattoe-Brown ^c and Peer-Olaf Siebers ^d

^aInstitute of Management Accounting and Simulation, Hamburg University of Technology, Germany; ^bNorwegian College of Fisheries Science, UiT The Arctic University of Norway, Norway; ^cSchool of Media, Communication and Sociology, University of Leicester, Leicester, UK; ^dSchool of Computer Science, University of Nottingham, Nottingham, UK

ABSTRACT

This article describes and justifies a reporting standard to improve data use documentation in Agent-Based Modelling. Following the development of reporting standards for models themselves, empirical modelling has now developed to the point where these standards need to take equally effective account of data use (which previously has tended to be an afterthought to model description). It is particularly important that a standard should allow the reporting of the *different* uses to which data may be put (specification, calibration and validation), but also that it should be compatible with the integration of different *kinds* of data (for example, survey, ethnographic and experimental) sometimes known as mixed methods research. The article motivates the need for standards generally, and positions the distinctive contribution of the RAT-RS reporting standard. It describes how the standard was developed to ensure its usability, presents and explains it, and describes possibilities for future development.

KEYWORDS

Agent-based modelling; data; interdisciplinarity; reporting standards; reproducibility

Motivation

The scientific method relies on progressive research. In order to build on previous work, we must both understand and (thus) trust it. Successful disciplines and methods do not only generate research, but (in the process) establish standards by which research should be judged. In mature fields, the latter process is often largely invisible. For example, broad agreement about how regression analysis is done and evaluated is transmitted by the normal socialisation of new researchers rather than by debate. As a newer field, Agent-Based Modelling must try to do research *and* establish standards in parallel. An Agent-Based Model, hereafter ABM, is a kind of computer simulation which explicitly represents ‘agents’ – aspects of social systems with *agency*, like individual decision makers or business organisations – in their interaction with each other and their environment. ABMs are indispensable in modelling *complex* systems: those in which relationships between individual and aggregate phenomena are not intuitive. For a more detailed discussion and example, see, Chattoe-Brown (2013). The present article is a contribution to the process of establishing standards in the area of data use for ABMs. Specifically, we want to establish standards which allow for the use of different kinds of data (sometimes for different purposes) within the same ABM. For example (but see, Chattoe-Brown, 2019 for more detail), data can be used in *specification* (deciding what elements an ABM should have i. e. interviewing people about how they make decisions so we know whether this should be modelled as rational, habitual or whatever), in

calibration (once we know that an ABM should involve friendship networks what numbers of friends do people typically have based on a survey?) and in *validation* (does simulated data resemble – according to some metric – the corresponding real data, for example, from an experiment?). This is particularly important since ABMs tend to focus on concrete phenomena (like unemployment or crime) rather than reproducing existing perspectives from distinct disciplines. This being so, any effective standard needs to be interdisciplinary.

The basic issue is very easily stated. As readers of Agent-Based Modelling research with many years of combined experience, the authors have often found that they cannot reconstruct how (and almost more importantly why) reported ABMs have used data as they have. This means that it is hard to evaluate such research effectively (and fairly), that it would be difficult to replicate it (for example) and, thus, that it cannot really be trusted as a basis for further scientific progress (Stodden et al., 2016). What can practically be done about this problem?

Previous research has already suggested an answer meeting with at least some academic approval. Devise a set of questions that researchers can relatively easily answer thus improving and standardising the reporting of data use in ABMs for their readers. Such a question set attempts to capture emerging best practice in a field, support self-discipline in committed researchers, and enable readers and reviewers better to evaluate research.

However, while the broad idea is straightforward, successful implementation (producing an effective standard that might be widely used) is not. Such question sets cannot simply reflect the views of a single discipline or method, but need to be genuinely mixed method and interdisciplinary. In particular the questions cannot assume – or even imply – for example, that only quantitative data matters. They must be based on recognised expertise and active participation in empirical ABM research rather than being philosophical exercises. They must be as simple and well designed as possible to be compatible with the goals of ensuring that they are firstly adopted and can then be used successfully.

This article thus proposes a question set constituting a reporting standard. Based on these questions, researcher responses should adequately document data use associated with a specific ABM for understanding, evaluation, and replication. We have named this the RAT-RS (Rigour And Transparency – Reporting Standard). As well as simply describing the standard and emphasising a mixed method understanding of data, the article shows how the RAT-RS should be used (based on an example) – and equally importantly *why* it should be used – and also explains how it was devised to maximise its value and usability.

Apart from its general thoroughness, practicability, and generality attained by repeated testing and contrasting to other existing standards, we see arguments from individual to community level for using the RAT-RS:

- support self-discipline for data reporting;
- structure the general thinking about data requirements;
- refine rationales for data use and create a shared procedural knowledge with the reader to enable better model assessments;
- facilitate establishing a common standard for data use to increase transparency and aid scientific progress in a maturing field.

The article is structured as follows. The next section provides a framework for understanding issues of data use and integration (particularly across disciplines and research methods) in Agent-Based Modelling. The third section describes other potentially relevant standards and shows how they justify the unique contribution of the RAT-RS. The fourth section explains how the RAT-RS was designed to achieve its goals. The fifth section describes practical issues in using the RAT-RS. The sixth section describes the RAT-RS itself and illustrates how it can be completed. The seventh section discusses testing the standard. The eighth section concludes and suggests further research and developments.

Challenges of data use and integration in Agent-Based Modelling

An ABM, like any other model, resembles a phenomenon that we wish to investigate. Typically, these phenomena are based on aspects we can observe in the real world. Data have the role of connecting the model and modelled world. In everyday language, data is the result of a measurement activity. This covers a broad range from data measured through surveys to data measured through human sensory channels. Therefore, data can be described as entities discernible from their context, defined by the specifics of measurement (Boisot & Canals, 2004; Floridi, 2015). Such data become information only if they have meaning, according to the widely recognised ‘General Definition of Information’ (Floridi, 2015).

As justified above, there is a need to clarify how data are used in ABMs and how they should be handled in modelling practice. As this themed section argues, an ABM plays an enabling role in ‘relating and integrating different kinds of evidence and types of knowledge’, which contributes to its increasing use in ‘inter- and transdisciplinary research as an integrating method’ (see the Introduction of this themed section). However, this flexibility creates a need for caution. Data themselves lack any meaning. Instead, meaning is given by the researcher who *uses* the data and by the context of the ABM. Therefore, their meaning might change when different data are combined, integrated or used for different purposes than that for which they were originally collected. For example, a survey validating one ABM might calibrate another.

Under perfect transparency, authors of an ABM would be able to convey each and every decision (and its underlying intentions) regarding how they progressed from their thoughts about the phenomena, to the collection and use of data and ultimately to the final design and evaluation of their model. Although full transparency (in this sense) is probably impossible, this issue nonetheless raises a severe challenge for Agent-Based Modelling as a research method. Other methods, like experiments or grounded theory, use more standardised procedures. Even though there are always ongoing debates about the right or best application of these methods, there is a broad shared understanding regarding how their data are used to generate results. Therefore, given the shared procedural knowledge of these methods, researchers also share common understandings about the meaning of their data.

This aim is more difficult to achieve for ABMs, which are characterised by a much more creative design process depending on individual modelling strategies and a constant iteration between phenomenon, data, conceptual model, and operational model. The model itself is the object of study (Weisberg, 2007). However, readers can only evaluate data use based on their knowledge of the published ABM and its description. This, unfortunately, often leaves questions unresolved about why data were used in certain ways. Thus (to avoid this), a better understanding of the model development process is necessary.

This leaves us with two particular challenges regarding data documentation for ABMs. First, the question of *why* data were used is not directly accessible through describing a model. Flexibility in the model development process also does not allow readers to derive such knowledge from shared understandings of standardised procedure. Hence, there is a crucial scientific responsibility to report *reasons* for particular data uses. Second, since the model development process generates a flood of decisions, often implicit but still leading to subsequent steps, a balance must be found regarding what to report. Maximal reporting is not a sensible solution and would probably not be adopted by modellers nor appreciated by their readers.

In the following section we analyse existing standards dealing with data documentation. We demonstrate that the above mentioned challenges are barely met and therefore, propose the RAT-RS as a solution.

Review of existing standards

In this section, we justify the distinctive contribution of the RAT-RS against relevant Agent-Based Modelling standards. Two important criteria for a general standard are generality and use. These criteria exclude several proposed approaches. If we define use (rather generously) as one citation

per year (including publication and present years) then we find the following (all citations in Web of Science, as of 24.06.2020). Wolf et al. (2013) – Dahlem ABM Documentation Guidelines – has been cited six times and is only intended for use with economic models. Parker et al. (2008) – MR POTATOHEAD – has been cited twice and is only intended to document Land Use Cover Change models. While Altaweel et al. (2010) – Delineate, Structure and Gather (DSG) – is meant to apply more generally, it has only been cited five times. Siebers and Klügl (2017) – Engineering Agent Based Social Simulation (EABSS) – is also of general application, but received three citations.

Having excluded approaches that are not widely used (or are not general), we now consider how RAT-RS makes a contribution distinct from those approaches with clear scholarly impact.

Interestingly, the most widely cited standards predominantly describe *models* rather than chiefly documenting *data* use, as the RAT-RS does. By far the best known is the ODD (Overview, Design concepts, and Details) protocol. Grimm et al. (2020) also references the original protocol, two updates, and a protocol variant adding decision making (with 2565 citations in total). In addition there is an ODD variant attending further to data (Reinhardt et al., 2018), but this is excluded here through insufficient citation. There is, thus, at least some reason to suppose that Agent-Based Modellers accept and are interested in reporting standards generally. The contribution of the RAT-RS can thus be characterised as documenting data use (and particularly its motivation) with as much detail and rigour as ODD documents models. However, despite its huge success, two issues with the ODD approach have motivated us to develop the RAT-RS independently. Firstly, it has taken a long time for ODD researchers to see the need for data and motivation in later variants of the protocol. We are thus concerned that the way these are handled may be distorted by the pre-existing protocol (which was designed without reference to data or motivation). We thus thought it useful to design a data standard directly from models using modellers to check that such path dependent distortion has not occurred. The second point is practical. Although the very latest version of the protocol discusses the value of rationale as part of an ODD, this is not illustrated as feasible practice. (The supplementary material in Grimm et al., 2020 actually provides illustrations of the – perhaps unfeasibly onerous – TRACE Protocol and not a rationale extension of ODD.) We thus intend that RAT-RS provides a parsimonious rationale specifically for data use and not for the entire model or modelling process.

Schmolke et al. (2010) – TRACE Protocol – has received 209 citations since publication, but is intended for ecological models and, thus, fails the generality test that RAT-RS (through the way it was constructed – see Methodology and Testing sections) aspires to pass. In addition, as with ODD, its (intermittent) references to data mainly serve *model* description and take a particular view of both data (quantitative) and its appropriate role (calibration rather than specification for example, – see, Chattoe-Brown, 2021).

Lorscheid et al. (2012) – Design of Experiments (DOE) – is a general approach cited 90 times since publication. However, despite the widespread perception that it contributes to data documentation, the article uses ‘data’ (perhaps unexpectedly) to mean almost exclusively *simulated* data and, thus, says nothing substantive about the model design process or validation (for example), which involves comparing empirical and simulated data.

Monks et al. (2019) – STrengthening the Reporting of Empirical Simulation Studies (STRESS) – has been cited 23 times and is, in a sense, intended for general use. However, it is designed to apply to Operations Research, which arguably has different goals than social simulation (solving assigned problems rather than understanding social behaviours). This is shown by the fact that empirically tested *description* does not really fit the discussion of model aims and that validation is explicitly excluded from the guidelines. Finally, we are concerned that the approach of synthesising existing guidelines (rather than at least partially testing documentation standards against real models as we did) may simply reproduce existing hazards for standard non-adoption (including implicit assumptions about data and the modelling process).

Richiardi et al. (2006) has been cited 96 times since publication. However, despite an excellent synthesis of issues, it arguably does not actually propose a documentation standard, but suggests how one might be established procedurally. This proposed process nonetheless informed our thinking in designing the RAT-RS.

Smajgl and Barreteau (2017) – Characterization and Parameterization (CAP) – has been cited six times since publication and makes a major investment in achieving effective generality (by applying its approach to 11 different case studies – an approach we also followed in designing the RAT-RS). It is also of unusual interest in contributing to the development of research design (i. e. systematic procedures for data use.) However, the article has a particular emphasis – what it describes as translating data into model assumptions – meaning, for example, that it has little to say about data in the context of validation. It is hard to be sure, but CAP may advocate model fitting – tuning parameters to ensure correspondence between real and simulated outputs – rather than independent validation. Despite its impressive contribution to modelling based on its own perspective, we believe that it cannot therefore serve as a *general* basis for documenting data use in ABMs. The non-equivalence (despite popular perception) of fitting and validation is discussed in Chattoe-Brown (2019).

Finally, Laatabi et al. (2018) – ODD+2D for Decision and Data – has been cited six times since publication. The fact that it is intended to have general application and to build on the ODD protocol (which is evidently a successful standard in use) makes this look the closest to the proposed RAT-RS. Because this contribution is so recent, it is not fair to expect it to be well cited already, but it should be noted that nobody except the authors have yet published an *application* of it. We believe (and our research agenda for RAT-RS echoes this) that authors being able to apply their own standard is an insufficient guide to general usefulness (which may explain why so many proposed standards haven't been widely adopted). Firstly, a single example that Laatabi et al. analyse may be untypically suited to their approach, therefore suffering selection bias. Further, while they may explain the standard to their own satisfaction, readers may not share their view that it is clear what to do or why it is worth doing it. Thus, we believe that deliberate user testing must be a key part of any successful standard.

Other concerns about the approach proposed by Laatabi et al. have also informed how we developed the RAT-RS, both conceptually and practically. Firstly, a standard is unlikely to be successful if it imposes a particular view of the modelling process and/or makes implicit assumptions about appropriate data. An important corollary of this is that the purpose of the RAT-RS is not to judge models from a single viewpoint – that others may not share – but to provide the information that allows it to be evaluated from a range of viewpoints. In a nutshell, our intended contrast is thus not between good and bad ABMs, but between ones that explain their data use adequately to a reader and those that don't. The example that Laatabi et al. analyse makes almost no mention of qualitative data and links quantitative data arbitrarily to calibration (rather than, for example, model specification). It will, thus, be of little use to those doing interviews or laboratory experiments or those wishing to ground other model aspects in data. In particular, Laatabi et al. neither make clear whether their ABM is validated (or how) or, from a methodological perspective, what research questions an ABM with calibration (but no validation) can answer. The RAT-RS is explicitly designed to handle all kinds of data, different possible uses for data (specification, calibration, and validation) and different approaches to modelling methodology. (See the Methodology section for more discussion of how this was achieved.)

Further, while we are keen not merely to duplicate the aspects of ODD dealing with motivation and research design, we find that many standards dealing with data suffer from 'atomisation' which impedes understanding. For example, Laatabi et al. provide a long list of variables and their possible states. But for the purposes of really engaging with their model, the reader mainly needs to know that householders take *multi-attribute* decisions (this crucial specification assumption does *not* seem to be justified empirically), whether the set of attributes modelled is empirically derived (if it is) or arbitrary, and how the actual proportions of dwellings with different attributes (and any empirically recorded correlations between these – we find no cheap houses that are very large, for example) were given empirical values. It has thus been our aim in RAT-RS to strike a careful balance between asking 'what' in regard to data and giving researchers the opportunity to provide sufficient 'why' to support the reader's understanding as well

(whether in terms of design decisions, methodology or practicalities as the case may be). Such an approach also makes the RAT-RS practically shorter since we have found that an ounce of rationale can save a pound of atomised description (which is potentially boring, both to write and read).

To sum up, based on existing contributions, our aim is an ABM reporting standard for data use meeting the following criteria:

- (1) It should be descriptive (not normative) and thus import no value judgements – whether explicitly or implicitly – about how modelling ought to be done (merely clarifying how it *was* done).
- (2) It should be explicitly designed for general use rather than applying to only one field or modelling approach. Nor should it originate from one field or approach, but, nonetheless, *claim* generality. It should not encapsulate (even tacitly) views that prioritise one kind of data, methodology (or ontology) or the perspective of one discipline. Its ease of use should not depend on the user sharing views with the designers. To this end, it should initially be developed and tested independently and not as an adjunct to existing standards which may distort its design with implicit preconceptions.
- (3) It should only require a minimum level of investment and effort from users relative to the length of a typical journal article. For example, the (incomplete) documentation shown by Altaweel et al. (2010) is only one page shorter than their article itself.
- (4) It should strike a sensible balance between motivating data use in ABMs and describing that data to avoid documentation that readers may find obscure and atomised.
- (5) Its specific focus on data documentation (rather than just seeing data as an adjunct to *model* documentation) would ideally make it *complementary* to successful protocols like ODD.

We now explain how these objectives were realised through the procedures followed to develop the RAT-RS.

Methodology

The designers of RAT-RS (who are also the article authors) come from a wide range of disciplinary backgrounds and approaches to modelling (management, natural resource governance, sociology, and computer science), thus ensuring that the RAT-RS was not captured by any one method or discipline. The initiative arose from a Lorentz Center workshop on Integrating Qualitative and Quantitative Evidence using Social Simulation (Leiden, the Netherlands, April 2019). At this workshop, we came together as a multi-disciplinary group of junior and senior modellers. Besides regular meetings of the designers, further steps in the RAT-RS development timeline were: Events at the Social Simulation Conference (Mainz, Germany, September 2019), the Second Workshop on Integrating Qualitative and Quantitative Evidence Using Social Simulation (Manchester, UK, November 2019) and the Third and Fifth Workshops in the same series (online, June and November 2020). These workshops were all organised by the European Social Simulation Association Special Interest Group ‘Using Qualitative Data to Inform Behavioural Rules’ <http://cfpm.org/qual2rule/>.

We started by developing an initial RAT-RS version using our joint experience. Soon we became aware of the need for variants of the RAT-RS, each considering a different driver for model development, and, after agreeing on a set of modelling approaches widely recognised in the literature, we created what we call individual RAT-RS *flavours*, based on our differing expertise. We presented one such flavour during a roundtable at Mainz (15 participants) and solicited feedback (testing activities, which took place in parallel are discussed in the Testing subsection below). At this stage, we also made the questions as brief and simple as possible. We also solicited feedback on presentations of progressively improved versions of the RAT-RS at the events in November (Manchester) and online in June (with a further 40 participants) and incorporated it into subsequent versions. To ensure that the formulated questions were reasonably applicable to subject areas targeted by the RAT-RS, we tested

them independently with several articles, some of our own and some selected from the literature. To avoid bias and arbitrariness, we finally passed the RAT-RS to 20 colleagues for a final feedback round, before writing this article, at the online event in November. Two participants also tested the RAT-RS by completing the questions for their own models. Another reason for testing the RAT-RS with a collection of articles was the common critique that a significant number of existing standards are based only on a single example chosen by like-minded researchers. We then harmonised the RAT-RS flavours by eliminating overlaps (questions that were the same in all flavours). What we provide here is therefore the first completed version of the RAT-RS to support reporting on interdisciplinary ABMs. We plan regular revisions to the reporting standard via feedback from its use.

Practicalities of using the RAT-RS

Before describing the RAT-RS questions in detail, this section presents guidance on its general use. Firstly, it is intended to be as intellectually flexible as possible. Although we recommend its completion at least partially *during* modelling (so all important decisions are recorded), we expect it will also be effective in summing up data use after research is completed (as we have shown with our examples) and, perhaps, in structuring general thinking about data requirements early on. To this end, the RAT-RS is also meant to be practically flexible. It is not necessary to complete it in strict order or complete all questions if they are not helpful. Instead, it is intended to be a structure that modellers can return to in an iterative way. Nonetheless, we recommend providing as much detail as possible and, where the user is not sure how to respond, that they follow their instinct and record their uncertainty in the ‘Any additional comments?’ section (which should also be used where the RAT-RS appears not to offer space for information the user thinks relevant. Access to this data by the RAT-RS designers should also progressively improve the scheme both in specific questions and overall structure.) The ‘Any additional comments?’ section can also be used to record developments during the modelling process that the user thinks remain relevant (for example, empirical approaches that appeared promising but were ultimately unfruitful.) Thirdly, the RAT-RS is designed not to interfere with other approaches, like TRACE and ODD. Nonetheless, we recommend it should be at least partially completed relatively early in the modelling process (and thus perhaps ahead of the bulk of TRACE recording and definitely before ODD, which its designers recommend completing after modelling.)

The RAT-RS

The next subsection discusses terminology used in the RAT-RS. As this is an interdisciplinary reporting standard, it is important to use a commonly acceptable vocabulary. Next, we show the questions constituting the RAT-RS and illustrative answers from a real example. Lastly, we describe the testing procedure we applied.

Terminology

Since the terminology of scientific modelling does not follow an overarching standard and is often contingent on communities and disciplines, we provide a contextualisation of key terms used in the RAT-RS.

In a scientific context, a model is an abstract representation of the system scientists analyse. Whether this system is hypothetical or real-world, although the latter is more common, the modeller must make certain preparations, among which describing a so-called target system is crucial. According to Weisberg (2007), target systems are described through a process of abstraction. In other words, the modeller decides which aspects of the system under investigation are relevant for the ABM, which itself depends on the phenomenon they are interested in investigating, research goal(s), and model purpose. In an idealised process, we might assume that a researcher would capture the target system they want to model by abstraction and then conceptualise it via model elements, based on a formal modelling logic.

Phenomena are ‘states or behaviours of a real-world system or group of systems’ (Elliott-Graves, 2020, p. 28) that are under scientific investigation. As such, a phenomenon can be examined from different angles and by different disciplines. Social inequality, for example, can be examined from a political or economic perspective. In practice the underlying research process can be messy, and it is usually characterised by an iteration between hypothetical/real-world phenomenon, target system, and model. If the conceptual model reaches a certain degree of maturity, the iterations also include the operational model and increasingly shift focus towards later stages of the modelling process. Nevertheless, specifying a target system makes particular sense under complexity where system elements and their inter-connections are not self-evident, and it is therefore difficult to identify exactly what produces the phenomenon. As Elliott-Graves (2020, p. 27) writes: ‘In fact, the value of specifying target-systems often becomes apparent only when scientists run into difficulties’.

The difference between the conceptual model and the target system is that the conceptual model is based on a pragmatic modelling perspective. This means that the tangible goal of creating a computable model is already manifested in the conceptual model which includes decisions about level of detail, assumptions and simplifications (Robinson, 2008). In other words, the conceptual model is seen through modelling glasses, while the target system is seen through more general glasses of scientific curiosity and the endeavour to understand and explain.

Questions

The RAT-RS consists of five question suites, labelled **model aim and context**, **conceptualisation**, **operationalisation**, **experimentation**, and **evaluation**. Together, these cover questions about data use for decision making, specification, calibration, output analysis, and validation within the full research life cycle of an ABM, i. e. from defining study purpose to evaluating experimental outputs. To account for different approaches that often drive modelling, we distinguish between theory driven, Operation Research (OR) data driven, model driven, and participatory RAT-RS flavours. This distinction only proves relevant for the conceptualisation question suite (and indeed we tried to achieve as much question standardisation as possible), as this is the only place within the RAT-RS where the basis for developing the ABM impacts the kinds of questions that need to be answered. For the other four question suites there is no difference between the flavours and therefore only a single set of questions is needed.

In the following discussion, we give brief overviews of the purposes for individual question suites, provide the questions themselves (which have explanations in braces where necessary), and show illustrative answers (in italics) for the model driven flavour. As stated in the Methodology section, we tested the RAT-RS for multiple cases, corresponding to the different flavours. In the response format for the questions, we allowed a choice between two approaches (based on the tester’s preference): Responding to each question using plain text or responding to some questions with plain text and others using predefined tables.

The full RAT-RS pack (user guide; questions suits for the different flavours; choice of test cases; set of templates) is available as a downloadable zip archive from <https://www.comses.net/codebases/f7e2c34a-4d07-4f37-9847-2b32df69528a/releases/1.0.0/>. For the base RAT-RS pack (user guide and question suits only), see Appendix A.

Question Suite 1 – Model Aim and Context Questions

Here the questions focus on the purpose and type (main driver) of the ABM, and help to define the target system. Possible type choices currently supported are theory driven models, which originate from a pre-existing theory or theories, OR data driven models, which focus on OR problems, model driven models, which originate from a pre-existing model or models, and participatory driven models, which focus on the use of participatory processes to design models. We use the term theory as a collective term for theory, theories, or theoretical constructs (something with any theoretical elements and not only what is understood as mature theory).

- (1) **If this RAT-RS use is related to a specific publication, please provide a reference to that publication.** *Chattoe-Brown, Edmund (2014) 'Using Agent Based Modelling to Integrate Data on Attitude Change', Sociological Research Online, 19(1), article 16, February. doi:0.5153/sro.3315*
- (2) **What is the purpose of the model?** {prediction, explanation, description, theoretical exploration, illustration, analogy, social interaction, or other (please specify); for an explanation of model purposes, see Edmonds et al. (2019)} *To show that data does not support a well-known model of attitude change (Zaller-Deffuant, hereafter ZD) and that another model can be devised which is somewhat compatible with data. (Edmonds et al., 2019: Explanatory use.)*
- (3) **What domain does the model research?** *Opinion dynamics.*
- (4) **What (research) question(s) is the model addressing?** {in general/in this publication} *Is the ZD model compatible with validation data? Is there a 'more realistic' model that is more compatible?*
- (5) **What is the MAIN driver for your initial model development step?** {theory(s), empirical evidence, existing model(s), participatory modelling data} *A model (ZD).*
- (6) **Explain why this MAIN driver was chosen?** *Because of concerns about the empirical status of widely cited models (of which ZD is an example).*
- (7) **What is the target system that this model reproduces?** {briefly describe the target system and its boundaries} *Random mixing of agents with opinions (and uncertainties) and the effect that interaction has on both of these individual attributes leading to macro outcomes (such as opinion polarisation). There is assumed to be no environment (such as advertising billboards) and no other agent classes (such as media outlets).*
- (8) **Explain why this target system and these boundaries were chosen.** *Because the model is widely cited. In fact, on reflection, the boundaries are obviously wrong (it is implausible there are no media effects at all), but that is not the issue that the study tackles.*

Any additional comments? *No.*

Question Suite 2 – Conceptualisation Questions: What and Why?

Here the questions help to provide information about how data was used for defining the 'What and Why?'. This refers to what aspects of the target system were considered for mapping into elements of the conceptual model and the empirical evidence supporting that decision. This question suite diverges for different RAT flavours because these aspects and evidence vary for distinct modelling approaches.

- (1) **What previous model is used (or models are used) as driver in this model? Give reference(s) to the model/models.** *This (ZD) model has many variants but two were used here: Deffuant, G., Amblard, F., Weisbuch, G. and Faure, T. (2002) 'How Can Extremism Prevail? A Study Based on the Relative Agreement Interaction Model', Journal of Artificial Societies and Social Simulation, 5(4)1, and Deffuant, G. (2006) 'Comparing Extremism Propagation Patterns in Continuous Opinion Models', Journal of Artificial Societies and Social Simulation, 9(3)8.*
- (2) **Why is/are this/these previous model(s) used?** *Example of a widely cited ABM with no empirical justification.*
- (3) **What are the elements of this/these previous model(s)?** *Agents, opinions, uncertainties, random mixing of agents and opinion/uncertainty updating via interaction.*
- (4) **Describe how you moved from the previous model elements to the elements of your model.** {distinguish (at least) between agents, environment, and relationships/interactions among any combination of these} *Based on 'social science common sense' and 'the literature' added several social processes to see if the resulting model would validate better.*

- (5) **Explain why elements of the previous model were included, excluded or changed in the current model.** *Uncertainty updating was too complicated for a model that did not fit the data anyway (so it was not implemented). The other elements were added to see if they improved validation. (Others could have been added or these could have been implemented differently however.) The aim was just ‘a’ validation (which ZD did not achieve at all) and not an empirically definitive ABM at this stage.*
- (6) **Explain why a model element was added when this was not included in the target system.** *Since the aim of the study was to show that validation was feasible (even though ZD did not attempt it) the elements were added to make successful validation more likely.*
- (7) **If an element of the previous model was changed, explain how it was done and why. Include sources if applicable.** *Added external media (the ‘stance’ of the paper you read can shift your opinion towards that of the paper and stance change is exogenous), static (or dynamic) social networks (not random mixing), social interaction does not always lead to convergence, but sometimes to divergence if opinions are very discrepant (the psychological literature is ambiguous about whether such repulsion effects exist – ZD assume they do not), multiple attitudes (rather than one) with movement towards (arbitrary) ‘ideologies’ (patterns in opinions) to reflect cognitive consistency as a phenomenon.*
- (8) **Describe the procedures and methods used to conceptualise the key target system elements as model elements. How did you make use of the evidence? What other sources did you utilise to conceptualise model elements?** *‘The literature’ implies there is a media effect, and the simplest process that could represent that was used (i. e. stances of papers changing exogenously). Where possible, public data was used (i. e. how many people read a daily paper and how many close friends people have). The possible divergence of opinions was based on some laboratory experiments. Correlations in multiple opinions were based (very loosely) on survey data analysis as a general effect.*

Any additional comments? No.

Question Suite 3 – Operationalisation Questions: How and Why?

Here the questions help to provide information about the How and Why of the ABM. This refers to how researchers considered what they said they would consider in the previous question suite, what data sources were used and what empirical evidence supports the decisions.

- (1) **What data element(s) did you include for implementing each key model element in the model’s scope?** *‘The literature’ for media effects, the roles of social networks, and correlations in multiple opinions. Survey data for newspaper readership and number of friends. Laboratory experiments for possible opinion divergence. Some elements were not really data based such as the dynamics of newspaper ‘stances’ (for example, going from pro to anti government).*
- (2) **Are these data elements implemented with the help of qualitative or quantitative data or further models?** *Quantitative data was derived from surveys and experiments. It is not really clear what kind of data ‘the literature’ is. No explicit qualitative data was used however.*
- (3) **Explain how data affected the way you implemented each model element and why. {i.e. explain your choice of data elements}** *Calibration where possible, but mainly the simplest model that might validate (while ZD does not).*
- (4) **What are the data elements used for in the modelling process: specification, calibration, validation, other?** *The distribution of opinions over time is validation (using British Social Attitude – BSA – data). All other data are used for calibration/specification.*

- (5) **Why for this use and not another one?** *The article tries to follow full Agent-Based Modelling methodology of Gilbert and Troitzsch (2005) so needs at least some independent data for both calibration and validation.*
- (6) **Did required data exist?** *Yes, except for the simplified social processes implemented i. e. what should the theory of the way that media change their positions on issues be?*
- (7) **If it existed, did you use it?** *Yes, as long as it was quick to find. (The point of the article was not to be right, but to show with how little effort you can be better as regards empirical modelling).*
- (8) **If you did not use it, why not?** *All data that could be found quickly were used.*
- (9) **For the existing data you used, provide details (a description) about data sources, sampling strategy, sample size, and collection period. For the data you collected, provide details about how it was collected, sampling strategy, sample size, and collection period.** *BSA survey (<https://www.bsa.natcen.ac.uk>). The questions used are given in the article. They were chosen to be controversial and to have good time series. It was reasoned that if attitudes on these matters do not polarise (which ZD seeks to explain) then none will.*
- (10) **Justify your data collection decisions from 3.9.** *The aim of the article justified these decisions. It was to show, although ZD did not calibrate or validate their model, that this was possible at quite low cost. Since the aim was not a true ABM, but only one that was somewhat validated, full data use was not necessary to the argument even if it had been possible (which, in the case of how newspapers change policy, for example, I suspect it is currently not).*
- (11) **If you needed to analyse or transform/manipulate the data before including them in the model (regardless if you collected data yourself or you used existing data), what did you do and why did you choose this specific approach?** *Simulation outputs were flattened (somewhat arbitrarily) into pro, neutral, and anti-opinions to match the BSA data categories.*
- (12) **In what format was the data implemented? {e.g. look-up table; distribution}** *Validation data were percentages in the three categories over time. Calibration/specification data were just single numbers or social processes.*
- (13) **Why this way?** *Simplicity/compatibility. Can a simple and easily populated empirical ABM nonetheless demonstrate validation? If so, how do we justify the continuing enthusiasm for non-empirical models?*

Any additional comments? *The study also validated the changes in opinion category size (i. e. from 27% pro to 12% pro) over the length of the 10 year real/simulated run.*

Question Suite 4 – Experimentation Questions

Here the questions help to provide information about calibration, experimental design, experimentation, and output analysis. In addition, they help to provide information about how the output was discussed with stakeholders (where applicable).

- (1) **Describe the calibration process you followed, stating which parameters you calibrated, their ranges, your reasons, and the similarity you achieved.** *Number of close friends (but network structure was arbitrary). Percentage of people reading newspapers (but dynamics of media ‘stance’ was arbitrary.) Replacing random mixing by networks was ‘common sense’ relaxation of an arbitrary restrictive assumption. (There is no evidence that social interaction is random.)*
- (2) **Describe the experimental design process you followed, stating your reasons, and the methods you used for the different steps. {e.g. calculating warm up period, run length, and number of replications; sensitivity analysis; robustness analysis}** *Single runs reported*

but multiple runs eyeballed for generality. Run in is not needed because actual opinion distributions are used at actual times. Validation was only eyeballing because ZD clearly fitted data so badly by comparison.

- (3) **What type(s) of experiments did you run? {e.g. calibration; empirical validation; sensitivity analysis; performance optimisation}** *Single runs.*
- (4) **For each experiment, name the purpose (objective).** *To show that the augmented model (with less unrealistic assumptions) validated better (at least had turning points) as the data do.*
- (5) **Describe the parameters you used to set up the experiments?** *Different initial opinion distributions given what the topic was. Different numbers of processes (i. e. with and without media effects) to see what combinations could be validated. No systematic exploration of process/parameter space however (not required for the purpose of the study).*
- (6) **Describe the data output that the model was designed to produce, your reasons for producing this output, and the data type of the output (qualitative or quantitative).** *Opinion distributions over time for ready docking to easily available data.*
- (7) **Describe the (statistical) analysis that you used on the output data and why.** *None. Just eyeballing. Did not need more rigour to make the point.*
- (8) **Did you discuss the output with the stakeholders? What did you discuss? Why? What effect did it have on the model?** *N/A.*

Any additional comments? *No.*

Question Suite 5 – Evaluation Questions

Here the questions help to provide information about the validation process and how validation results were discussed with stakeholders (where applicable).

- (1) **In validation, what similarity measures did you use and why? What similarity did you get? What would you consider a good similarity and why?** *None. Just eyeballing. Did not need more rigour to make the point.*
- (2) **How do the data outputs support an answer to the research question?** *ZD outputs look nothing like the real data in fact. Outputs from the improved model do somewhat. Therefore validation is not an impossible goal requiring epic data analysis and should therefore be more strongly required in Agent-Based Modelling.*
- (3) **Did you discuss the validation results with the participants? What did you discuss? Why? What effect did it have on the conclusions?** *N/A.*

Any additional comments? *No.*

Testing

When evaluating the effectiveness of the RAT-RS, we adopted the three qualitative criteria from Smajgl and Barreteau (2017): genericity and, hence applicability across diverse cases; capacity to effectively structure the process of reporting data use and integration; capacity to lead modellers to report all details necessary to allow for understanding, evaluation, and replication.

The standard was tested several times by the authors during the design process (see the Methodology section for more details). Independently from the authors, the standard was further tested on two occasions. A subset of questions was tested with 15 volunteers at the Mainz round-table. They each completed one flavour based on their own ABMs. A full version of the standard was then tested at the 5th Workshop on Integrating Qualitative and Quantitative Evidence Using Social Simulation.

The 17 modellers who tested the RAT-RS came from diverse communities (e. g. social science, natural resource management, economics, health sciences) and were able to relate the questions to their own data use in building an ABM. This suggests that the standard is sufficiently generic. Regarding the second criterion, we conclude that the standard captures all relevant steps in the process of data use and integration when building an ABM. As such, no respondent ever indicated that the standard disregarded any use of data in Agent-Based Modelling as the method is actually practiced, again suggesting that it is comprehensive and sufficiently generic. Regarding the third criterion, discussion of the examples suggests that RAT-RS helps modellers to report all the necessary detail that allows readers to understand and evaluate data use decisions.

Conclusion

The aim of this article has been to develop and justify a reporting standard that supports the effective documentation of data use and integration in Agent-Based Modelling and present the so-called RAT-RS at a particular stage in its development (as part of the plan to promote its use). We have shown why such a standard is needed, how we have designed it to be as easy to adopt and use as possible, what it consists of and how to use it. Based on testing so far, we have grounds for believing that using the RAT-RS will improve the rigour and transparency of data use in ABMs and thus improve the quality of research, reviewing, and replication. To facilitate its use, we believe we have demonstrated that the standard is general, intuitive, and concise (the example in the Questions subsection was completed in 1200 words).

But obviously, at the end of the day, it is users and not designers who must say whether something is usable or not. To this end, a large element of further research on the standard will be user testing. We want to ensure that the standard really is as general and intuitive as we currently believe it to be (based on the development process described above). To this end, we plan to design and implement a series of activities to promote user testing. These will include workshops where modellers complete and discuss their experience of the RAT-RS, cooperation with existing institutions (where RAT-RS completion can be part of a summer school or training course on Agent-Based Modelling), and ensuring that all the raw materials needed to complete the RAT-RS standard are adequately promoted, freely available, and well documented (and that the authors can continue to be contacted by potential and actual users.) We also recognise that adoption of innovations has a psychological element and that we may need to manage expectations convincingly (for example about how time consuming the RAT-RS is to complete) and use accessible dissemination strategies (like YouTube tutorials) as well as more traditional scholarly outlets.

More generally, this article goes about as far as it can based on the insight of the authors and informal respondents alone. While the RAT-RS can undoubtedly be improved (and will turn out to have limitations) the best way to deal with this is by promoting its use and responding to the issues that arise. This is also a tradition in the field of documentation standards where Grimm et al. regularly publish updates to the ODD protocol.

Nonetheless, there are areas where we can already envisage productive further development. The first of these is that, while we are reasonably confident that the RAT-RS flavours cover the commonest approaches to modelling, it may well be that new flavours or combinations will be needed based on user feedback. Secondly, if the RAT-RS proves to be successful, we suspect that an online version will facilitate both use of the reporting standard and the collation of data to drive further improvements. Thirdly, we will investigate if and how RAT-RS can best be used with other documentation tools, such as the ODD protocol.

In methodological communities where procedural standards are either not achievable or not yet achieved, other media for safeguarding transparency and rigor are needed. An open and ongoing discussion of such topics is therefore particularly needed and can be very fruitful. This includes appreciating and learning from previously failed and currently successful attempts as well as

a continuous development and periodic reflection on and updating of established practices. We welcome colleagues who would like to join these discussions and testing activities. The RAT-RS is designed not just for reporting on data use, but the process of completing it hopefully changes how people think about their use of data and its justification.

Acknowledgement

Edmund Chattoe-Brown's analysis was funded by the project "Towards Realistic Computational Models Of Social Influence Dynamics" (ES/S015159/1) funded by ESRC via ORA Round 5 (PI: Professor Bruce Edmonds, Centre for Policy Modelling, Manchester Metropolitan University: <https://gtr.ukri.org/projects?ref=ES%2FS015159%2F1>).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Sebastian Achter is a PhD candidate at the Institute of Management Accounting and Simulation (MAccS) at Hamburg University of Technology. He holds a Master's degree in International Management and Engineering. His research interests are the human factor in organizational forecasting practices and management control issues more broadly. While aspiring methodical pluralism he is a proponent of computational simulation methods, specifically agent-based simulation, as an integrative tool to investigate complex social phenomena.

Melania Borit is an interdisciplinary researcher with a passion for knowledge integration. She has a wide range of interconnected research interests: social simulation, agent-based modelling; research methodology; Artificial Intelligence ethics; pedagogy and didactics in higher education, games and game-based learning; culture and fisheries management, seafood traceability; critical futures studies. She is the leader of the CRAFT Lab – Knowledge Integration and Blue Futures at UiT The Arctic University of Norway.

Edmund Chattoe-Brown's career has been interdisciplinary (including Politics, Philosophy, Economics, Artificial Intelligence, Medicine, Law and Anthropology), focusing on the value of research methods (particularly Agent-Based Modelling) in generating warranted social knowledge. His aim has been to make models both more usable generally and particularly more empirical (because the most rigorous social scientists tend to be empirical.) The results of his interests have been published in 17 different peer reviewed journals across the sciences to date.

Peer-Olaf Siebers is an Assistant Professor at the School of Computer Science, University of Nottingham, UK. His main research interest is the application of Computer Simulation and Artificial Intelligence to study human-centric and coupled human-natural systems. He is a strong advocate of Object-Oriented Agent-Based Social Simulation and is advancing the methodological foundations. It is a novel and highly interdisciplinary research field, involving disciplines like Social Science, Economics, Psychology, Geography, Operations Research, and Computer Science.

ORCID

Sebastian Achter  <http://orcid.org/0000-0003-1755-4010>

Melania Borit  <http://orcid.org/0000-0002-1305-8581>

Edmund Chattoe-Brown  <http://orcid.org/0000-0001-8232-6896>

Peer-Olaf Siebers  <http://orcid.org/0000-0002-0603-5904>

References

- Altaweel, M. R., Alessa, L. N., Kliskey, A., & Bone, C. (2010). A framework to structure agent-based modeling data for social-ecological systems. *Structure and Dynamics*, 4(1). <https://escholarship.org/uc/item/7kw1h48n>.
- Boisot, M., & Canals, A. (2004). Data, Information and Knowledge: Have We Got It Right? *Journal of Evolutionary Economics*, 14(1), 43–67. <https://doi.org/10.1007/s00191-003-0181-9>
- Chattoe-Brown, E. (2013). Why Sociology Should Use Agent Based Modelling. *Sociological Research Online*, 18(3), 31–41. <https://doi.org/https://doi.org/10.5153/sro.3055>
- Chattoe-Brown, E. (2019). Agent based models. In P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, & R. A. Williams (Eds.), *SAGE Research Methods Foundations*. <https://doi.org/10.4135/9781526421036836969>

- Chattoe-Brown, E. (2021). Why Questions Like “Do Networks Matter?” Matter to Methodology: How Agent-Based Modelling Makes It Possible to Answer Them. *International Journal of Social Research Methodology*, 24(4), 429–442. <https://doi.org/10.1080/13645579.2020.1801602>
- Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañaola-Sales, C., Ormerod, P., Root, H., & Squazzoni, F. (2019). Different modelling purposes. *Journal of Artificial Societies and Social Simulation*, 22(3). <https://doi.org/10.18564/jasss.3993>
- Elliott-Graves, A. (2020). What is a Target System? *Biology & Philosophy*, 35(2), 2. <https://doi.org/10.1007/s10539-020-09745-3>
- Floridi, L. (2015). Semantic conceptions of information. In E. N. Zalta Ed. *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Laboratory, Stanford University. 07.01.15revision, accessed via on 29. 01.21 <https://plato.stanford.edu/entries/information-semantic/>
- Gilbert, N., & Troitzsch, K. G. (2005). *Simulation for the social scientist* (second ed.). Open University Press.
- Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S. A., Milles, A., Nabe-Nielsen, J., Polhill, J. G., Radchuk, V., Rohwäder, M.-S., Stillman, R. A., Thiele, J. C., & Ayllón, D. (2020). The ODD protocol for describing agent-based and other simulation models: A second update to improve clarity, replication, and structural realism. *Journal of Artificial Societies and Social Simulation*, 23(2). <https://doi.org/10.18564/jasss.4259>
- Laatabi, A., Marilleau, N., Nguyen-Huu, T., Hbid, H., & Babram, M. A. (2018). ODD+2D: an ODD based protocol for mapping data to empirical ABMs. *Journal of Artificial Societies and Social Simulation*, 21(2). <https://doi.org/10.18564/jasss.3646>
- Lorscheid, I., Heine, B.-O., & Meyer, M. (2012). Opening the ‘black box’ of simulations: Increased transparency and effective communication through the systematic design of experiments. *Computational and Mathematical Organization Theory*, 18(1), 22–62. <https://doi.org/10.1007/s10588-011-9097-3>
- Monks, T., Currie, C. S. M., Onggo, B. S., Robinson, S., Kunc, M., & Taylor, S. J. E. (2019). Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. *Journal of Simulation*, 13(1), 55–67. <https://doi.org/10.1080/17477778.2018.1442155>
- Parker, D. C., Brown, D., Polhill, J. G., Manson, S., & Deadman, P., (2008). *Illustrating a new “conceptual design pattern” for agent-based models and land use via five case studies: The MR POTATOHEAD framework*, A. L. Paredes & C. H. Iglesias eds., Agent Based Modelling in Natural Resource Management INSISOC.
- Reinhardt, O., Rucheinski, A., & Uhrmacher, A. M. (2018) ‘ODD+P: Complementing the ODD Protocol with Provenance Information’, in M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, & B. Johansson (eds.) *Proceedings of the 2018 Winter Simulation Conference (WSC)*, Gothenburg, Sweden, 2018, pp. 727–738, <https://doi.org/10.1109/WSC.2018.8632481>
- Richiardi, M., Leombruni, R., Saam, N., & Sonnessa, M. (2006). A common protocol for agent-based social simulation. *Journal of Artificial Societies and Social Simulation*, 9(1). <<http://jasss.soc.surrey.ac.uk/9/1/15.html>>
- Robinson, S. (2008). Conceptual Modelling for Simulation Part I: Definition and Requirements. *Journal of the Operational Research Society*, 59(3), 278–290. <https://doi.org/10.1057/palgrave.jors.2602368>
- Schmolke, A., Thorbek, P., DeAngelis, D. L., & Grimm, V. (2010). Ecological models supporting environmental decision making: A strategy for the future. *Trends in Ecology & Evolution*, 25(8), 479–486. <https://doi.org/10.1016/j.tree.2010.05.001>
- Siebers, P.-O., & Klügl, F. (2017). What software engineering has to offer to agent-based social simulation. In B. Edmonds & R. Meyer (Eds.), *Simulating social complexity, understanding complex systems series* (second ed., pp. 81–117). Springer-Verlag.
- Smaijl, A., & Barreteau, O. (2017). Framing options for characterising and parameterising human agents in empirical ABM. *Environmental Modelling and Software*, 93, 29–41. <https://doi.org/10.1016/j.envsoft.2017.02.011>
- Stodden, V., McNutt, M., Bailey, D. H., Deelman, E., Gil, Y., Hanson, B., Heroux, M. A., Ionnidis, J. P. A., & Taufer, M. (2016). Enhancing reproducibility for computational methods. *Science*, 354(6317), 1240–1241. <https://doi.org/10.1126/science.aah6168>
- Weisberg, M. (2007). Who is a modeler? *The British Journal for the Philosophy of Science*, 58(2), 207–233. <https://doi.org/10.1093/bjps/axm011>
- Wolf, S., Bouchaud, J.-P., Cecconi, F., Cincotti, S., Dawid, H., Gintis, H., van der Hoog, S., Jaeger, C. C., Kovalevsky, D. V., Mandel, A., & Paroussos, L. (2013). Describing economic agent-based models: Dahlem ABM documentation guidelines. *Complexity Economics*, 2(1), 63–74. accessed at on 21.10.21 https://www.researchgate.net/profile/Sander-Van-Der-Hoog/publication/257558097_Describing_economic_agent-based_models_-_Dahlem_ABM_documentation_guidelines/links/5a3a94534585155ac76f56d2/Describing-economic-agent-based-models-Dahlem-ABM-documentation-guidelines.pdf

Appendix A: RAT-RS Base Version Pack v1.0

Please note that what we are presenting here is the ‘Base Version’ of the RAT-RS Pack (excluding alternative formats and examples) using the original stylesheet. The ‘Full Version’ of the RAT-RS Pack can be found at <https://www.comses.net/codebases/f7e2c34a-4d07-4f37-9847-2b32df69528a/releases/1.0.0/>.

- A1: RAT-RS Base Version Pack User Guide v1.0
- A2: Theory Driven RAT-RS Question Suite v1.0
- A3: OR Data Driven RAT-RS Question Suite v1.0
- A4: Another Model Driven RAT-RS Question Suite v1.0
- A5: Participatory Driven RAT-RS Question Suite v1.0

A1: RAT-RS Base Version Pack User Guide v1.0

Last update: 20/10/2021

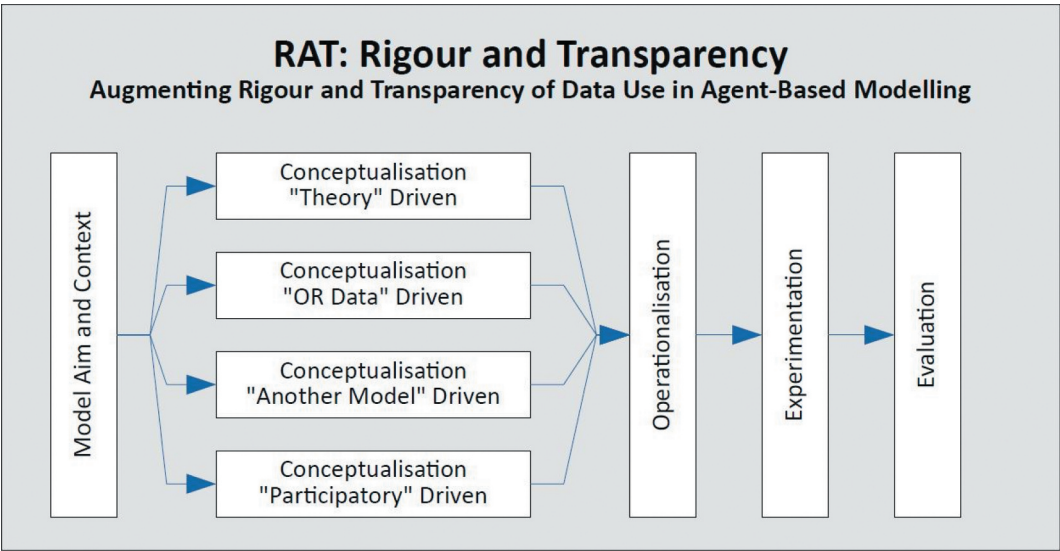
RAT-RS Developers

- Sebastian Achter, Hamburg University of Technology, Germany
- Melania Borit, UiT The Arctic University of Norway, Norway
- Edmund Chattoe-Brown, University of Leicester, UK
- Peer-Olaf Siebers, University of Nottingham, UK

Summary

The Rigor and Transparency Reporting Standard (RAT-RS) is a tool to improve the documentation of data use in Agent-Based Modelling. Following the development of reporting standards for models themselves, attention to empirical models has now reached a stage where these standards need to take equally effective account of data use (which until now has tended to be an afterthought to model description). It is particularly important that a standard should allow the reporting of the different uses to which data may be put (specification, calibration and validation), but also that it should be compatible with the integration of different kinds of data (for example, statistical, qualitative, ethnographic and experimental) sometimes known as mixed methods research. For the full details on the RAT-RS, please refer to the related publication ‘RAT-RS: A Reporting Standard for Improving the Documentation of Data Use in Agent-Based Modelling’ (Achter et al. under review).

The following figure shows the RAT-RS question suites, where the box names indicate the focus of a specific question suite. The RAT-RS comes in ‘flavours’ with distinct Conceptualisation question suites (horizontal boxes). So, the first step when aiming to use the RAT-RS is to identify the MAIN driver for the initial model development step. Possible choices currently supported by the RAT-RS are ‘theory-driven’ models, which focus on a pre-existing theory or theories, ‘OR-data-driven’ models, which focus on key mechanisms, ‘another-model-driven’ models, which focus on a pre-existing model or models, and ‘participatory-driven’ models, which focus on the use of participatory processes to design models.



There are two different versions of the RAT-RS Pack, the ‘Base Version’ (which comes with all questions, but excludes alternative formats and examples) and the ‘Full Version’ (which includes the ‘Base Version’ and alternative formats and examples). While the ‘Base Version’ is supplementary material for our journal article (Achter et al., under review), the ‘Full Version’ can be downloaded from the following resource: <https://www.comses.net/codebases/f7e2c34a-4d07-4f37-9847-2b32df69528a/releases/1.0.0/>.

Note that not all questions have to be answered. However, you are encouraged to provide as much information as possible. Remember that this is a documentation tool. It is for capturing data use and its reasons in a non-judgemental way. If a question is unclear to you, please answer it following your instinct, and leave a remark in the ‘Any additional comments?’ space of the related section. If you want to capture information about changes you made to the use of data over time (i.e. if you changed your mind while modelling), you can also provide this information together with some justification in the ‘Any additional comments?’ space of the related section.

Since the used terminology in scientific modelling does not follow an overarching standard and is often contingent on communities and disciplines, we provide a contextualisation of key terms that are used in the RAT-RS. You can find our current ‘Glossary of Terms’ at the end of this user guide.

A final note. If you want to help us to develop this RAT-RS further, please email us a copy of your filled in RAT-RS document(s) and share your experience with filling it in. We also welcome suggestions for future improvements. Emails can be sent to peer-olaf.siebers@nottingham.ac.uk.

Reference

- Achter S, Borit M, Chattoe-Brown E, and Siebers PO (2022) “RAT-RS: A Reporting Standard for Improving the Documentation of Data Use in Agent-Based Modelling”, International Journal of Social Research Methodology, <http://dx.doi.org/10.1080/13645579.2022.2049511>.

Glossary of Terms

Term	Definition	Evidence
Data	Data can be described as entities discernible from their context, defined by the specifics of measurement (Boisot & Canals, 2004; Floridi, 2015). Such data become information only if they have meaning according to the widely recognised “General Definition of Information” (Floridi, 2015).	Boisot, M. & Canals, A. (2004). “Data, information and knowledge: have we got it right?” Journal of Evolutionary Economics, 14(1), 43-67. Floridi (2015). “Semantic Conceptions of Information”. In: The Stanford Encyclopedia of Philosophy (Winter 2016 Edition). https://plato.stanford.edu/archives/ win2016/entries/informa tion-semantic/ .

(Continued)

Term	Definition	Evidence
Data element	A data element is data corresponding to a model element. So, for the disease example, the data element for contact might be a survey (or diary) of social contacts while for disease progression it might be medical histories/records (and perhaps even laboratory data on how the disease 'works').	Our own definition.
Domain	A distinct area of research interest independent of disciplines or applied methods.	Our own definition.
Driver	Main starting point for the development of the agent-based model	Our own definition.
Empirical evidence	Scientific evidence produced by empirical research methods.	Our own definition.
Model	An abstract representation of reality (e.g. a phenomenon, elements of the real world, or a real world system of interacting elements)	Partly from the ResearchGate answer to the question, 'A model is usually believed to be an abstract representation of reality. How abstract should a specification be to be considered a model?'. Examples are our own.
Model element	A model element is part of a model that may be somewhat self-contained and/or distinct in its operation. For example, in a disease model you need both a model element for the contact process (who meets who and where) and another for the disease progression process (what happens once you are infected). One refers to an individual and their physiology and is thus somewhat distinct from the other, which refers to interactions in the social world.	Our own definition.
Phenomenon	Phenomena are states or behaviours of a real-world system or group of systems, studied by a particular discipline, such as population growth, competition, or predation.	Elliott-Graves (2020). "What is a target system?" <i>Journal of Biology & Philosophy</i> , 35(2), 1-22 (p. 28).
Qualitative vs quantitative data	Qualitative data is defined as non-numerical data, such as text, video, photographs or audio recordings. Therefore, quantitative data are defined as numerical, such as closed survey responses or sensor data.	https://www.simplypsychology.org/qualitative-quantitative.html with our own additions.
Scientific evidence	The available body of facts or information indicating whether a belief or proposition is true or valid.	https://www.lexico.com/definition/evidence
Target system	A target system is those aspects of the real-world system that are studied in order to gain knowledge about the phenomenon.	Elliott-Graves (2020). "What is a target system?" <i>Journal of Biology & Philosophy</i> 35(2) 1-22 (p. 28).
Theory	Collective term for theory, theories, or theoretical constructs; generally something theoretical and not only what is understood as a 'mature' theory	Our own definition.
Theoretical vs empirical	Theoretical means originating from abstract theory while empirical means originating from empirical data.	Our own definition.

A2: Theory Driven RAT-RS Question Suite v1.0

Filled in by (name; email): _____

Signed off on (date): _____

1	MODEL AIM AND CONTEXT QUESTIONS
1.1	If this RAT-RS use is related to a specific publication, please provide a reference to that publication.
1.2	What is the purpose of the model? {prediction, explanation, description, theoretical exploration, illustration, analogy, social interaction, or other (please specify); for an explanation of model purposes, see, Edmonds et al. (2019) http://jasss.soc.surrey.ac.uk/22/3/6.html }
1.3	What domain does the model research?
1.4	What (research) question(s) is the model addressing? {in general/in this publication}
1.5	What is the MAIN driver for your initial model development step? {theory(s), empirical evidence, existing model(s), participatory modelling data}
1.6	Explain why this MAIN driver was chosen?
1.7	What is the target system that this model reproduces? {briefly describe the target system and its boundaries}
1.8	Explain why this target system and these boundaries were chosen.
1AQ	Any additional comments?

2	CONCEPTUALISATION QUESTIONS: WHAT AND WHY?
2.1	What theory is used (or theories are used) as driver in this model? Give reference(s) to the theory/theories.
2.2	Why is/are this/these theory/theories used?
2.3	What are the elements of the theory?
2.4	Which of these theory elements were mapped into model elements? {distinguish (at least) between agents, environment, and relationships/interactions among any combination of these}
2.5	Explain why theory elements were included, excluded or changed in the model.
2.6	Explain why a model element was added when this was not included in the target system
2.7	If a theory element was (or theory elements were) changed, explain how it was done and why. Include sources if applicable.
2.8	Describe the procedures and methods used to conceptualise the target system elements as model elements. How did you make use of the evidence? What other sources did you utilise to conceptualise model elements?
2.9	N/A.
2AQ	Any additional comments?

3	OPERATIONALISATION QUESTIONS: HOW AND WHY?
3.1	What data element(s) did you include for implementing each key model element in the model's scope?
3.2	Are these data elements implemented with the help of qualitative or quantitative data or further models?
3.3	Explain how data affected the way you implemented each model element and why. {i.e. explain your choice of data elements}
3.4	What are the data elements used for in the modelling process: specification, calibration, validation, other?
3.5	Why for this use and not another one?
3.6	Did required data exist?
3.7	If it existed, did you use it?
3.8	If you did not use it, why not?
3.9	For the existing data you used, provide details (a description) about data sources, sampling strategy, sample size, and collection period. For the data you collected, provide details about how it was collected, sampling strategy, sample size, and collection period.
3.10	Justify your data gathering decisions from 3.9.
3.11	If you needed to analyse the data before including them in the model (regardless if you collected data yourself or you used existing data), what data analysis did you do and why did you choose this specific analysis?
3.12	In what format was the data implemented? {e.g. look-up table; distribution}
3.13	Why this way?
3AQ	Any additional comments?

4	EXPERIMENTATION QUESTIONS
4.1	Describe the calibration process you followed, stating which parameters you calibrated, their ranges, your reasons, and the similarity you achieved.
4.2	Describe the experimental design process you followed, stating your reasons, and the methods you used for the different steps. {e.g. calculating warm up period, run length, and number of replications; sensitivity analysis; robustness analysis}
4.3	What type(s) of experiments did you run? {e.g. calibration; empirical validation; sensitivity analysis; performance optimisation}
4.4	For each experiment, name the purpose (objective).
4.5	Describe the parameters you used to set up the experiments?
4.6	Describe the data output that the model was designed to produce, your reasons for producing this output, and the data type of the output (qualitative or quantitative).
4.7	Describe the (statistical) analysis that you used on the output data and why.
4.8	Did you discuss the output with the stakeholders? What did you discuss? Why? What effect did it have on the model?
4AQ	Any additional comments?

5	EVALUATION QUESTIONS
5.1	In validation, what similarity measures did you use and why? What similarity did you get? What would you consider a good similarity and why?
5.2	How do the data outputs support an answer to the research question?
5.3	Did you discuss the validation results with the participants? What did you discuss? Why? What effect did it have on the conclusions?
5AQ	Any additional comments?

A3: OR Data Driven RAT-RS Question Suite v1.0

Filled in by (name; email): _____

Signed off on (date): _____

1	MODEL AIM AND CONTEXT QUESTIONS
1.1	If this RAT-RS use is related to a specific publication, please provide a reference to that publication.
1.2	What is the purpose of the model? {prediction, explanation, description, theoretical exploration, illustration, analogy, social interaction, or other (please specify); for an explanation of model purposes, see, Edmonds et al. (2019) http://jasss.soc.surrey.ac.uk/22/3/6.html }
1.3	What domain does the model research?
1.4	What (research) question(s) is the model addressing? {in general/in this publication}
1.5	What is the MAIN driver for your initial model development step? {theory(s), empirical evidence, existing model(s), participatory modelling data}
1.6	Explain why this MAIN driver was chosen?
1.7	What is the target system that this model reproduces? {briefly describe the target system and its boundaries}
1.8	Explain why this target system and these boundaries were chosen.
1AQ	Any additional comments?

2	CONCEPTUALISATION QUESTIONS: WHAT AND WHY?
2.1	What key mechanism(s) of the target system is/are used as driver(s) in this model? {examples: queuing system; physical/ social/political network}
2.2	Why is/are this/these key mechanism(s) used?
2.3	What are the key elements of the target system?
2.4	Which of these key elements were mapped into model elements? {distinguish (at least) between agents, environment, and relationships/interactions among any combination of these}
2.5	Explain why key elements of the target system were included, excluded or changed in the model.
2.6	Explain why a model element was added when this was not included in the target system
2.7	Describe the procedures and methods used to conceptualise the target system key elements as model elements. How did you make use of the evidence? What other sources did you utilise to conceptualise model elements?
2.8	N/A.
2.9	N/A.
2AQ	Any additional comments?

3	OPERATIONALISATION QUESTIONS: HOW AND WHY?
3.1	What data element(s) did you include for implementing each key model element in the model's scope?
3.2	Are these data elements implemented with the help of qualitative or quantitative data or further models?
3.3	Explain how data affected the way you implemented each model element and why. (i.e. explain your choice of data elements)
3.4	What are the data elements used for in the modelling process: specification, calibration, validation, other?
3.5	Why for this use and not another one?
3.6	Did required data exist?
3.7	If it existed, did you use it?
3.8	If you did not use it, why not?
3.9	For the existing data you used, provide details (a description) about data sources, sampling strategy, sample size, and collection period. For the data you collected, provide details about how it was collected, sampling strategy, sample size, and collection period.
3.10	Justify your data gathering decisions from 3.9.
3.11	If you needed to analyse the data before including them in the model (regardless if you collected data yourself or you used existing data), what data analysis did you do and why did you choose this specific analysis?
3.12	In what format was the data implemented? (e.g. look-up table; distribution)
3.13	Why this way?
3AQ	Any additional comments?

4	EXPERIMENTATION QUESTIONS
4.1	Describe the calibration process you followed, stating which parameters you calibrated, their ranges, your reasons, and the similarity you achieved.
4.2	Describe the experimental design process you followed, stating your reasons, and the methods you used for the different steps. (e.g. calculating warm up period, run length, and number of replications; sensitivity analysis; robustness analysis)
4.3	What type(s) of experiments did you run? (e.g. calibration; empirical validation; sensitivity analysis; performance optimisation)
4.4	For each experiment, name the purpose (objective).
4.5	Describe the parameters you used to set up the experiments?
4.6	Describe the data output that the model was designed to produce, your reasons for producing this output, and the data type of the output (qualitative or quantitative).
4.7	Describe the (statistical) analysis that you used on the output data and why.
4.8	Did you discuss the output with the stakeholders? What did you discuss? Why? What effect did it have on the model?
4AQ	Any additional comments?

5	EVALUATION QUESTIONS
5.1	In validation, what similarity measures did you use and why? What similarity did you get? What would you consider a good similarity and why?
5.2	How do the data outputs support an answer to the research question?
5.3	Did you discuss the validation results with the participants? What did you discuss? Why? What effect did it have on the conclusions?
5AQ	Any additional comments?

A4: Another Model Driven RAT-RS Question Suite v1.0

Filled in by (name; email): _____

Signed off on (date): _____

1	MODEL AIM AND CONTEXT QUESTIONS
1.1	If this RAT-RS use is related to a specific publication, please provide a reference to that publication.
1.2	What is the purpose of the model? (prediction, explanation, description, theoretical exploration, illustration, analogy, social interaction, or other (please specify); for an explanation of model purposes, see, Edmonds et al. (2019) http://jasss.soc.surrey.ac.uk/22/3/6.html);
1.3	What domain does the model research?

(Continued)

1	MODEL AIM AND CONTEXT QUESTIONS
1.4	What (research) question(s) is the model addressing? {in general/in this publication}
1.5	What is the MAIN driver for your initial model development step? {theory(s), empirical evidence, existing model(s), participatory modelling data}
1.6	Explain why this MAIN driver was chosen?
1.7	What is the target system that this model reproduces? {briefly describe the target system and its boundaries}
1.8	Explain why this target system and these boundaries were chosen.
1AQ	Any additional comments?
2	CONCEPTUALISATION QUESTIONS: WHAT AND WHY?
2.1	What previous model is used (or models are used) as driver in this model? Give reference(s) to the model/models.
2.2	Why is/are this/these previous model(s) used?
2.3	What are the elements of this/these previous model(s)?
2.4	Describe how you moved from the previous model elements to the elements of your model. {distinguish (at least) between agents, environment, and relationships/interactions among any combination of these}
2.5	Explain why elements of the previous model were included, excluded or changed in the current model.
2.6	Explain why a model element was added when this was not included in the target system.
2.7	If an element of the previous model was changed, explain how it was done and why. Include sources if applicable.
2.8	Describe the procedures and methods used to conceptualise the key target system elements as model elements. How did you make use of the evidence? What other sources did you utilise to conceptualise model elements?
2.9	N/A.
2AQ	Any additional comments?
3	OPERATIONALISATION QUESTIONS: HOW AND WHY?
3.1	What data element(s) did you include for implementing each key model element in the model's scope?
3.2	Are these data elements implemented with the help of qualitative or quantitative data or further models?
3.3	Explain how data affected the way you implemented each model element and why. {i.e. explain your choice of data elements}
3.4	What are the data elements used for in the modelling process: specification, calibration, validation, other?
3.5	Why for this use and not another one?
3.6	Did required data exist?
3.7	If it existed, did you use it?
3.8	If you did not use it, why not?
3.9	For the existing data you used, provide details (a description) about data sources, sampling strategy, sample size, and collection period. For the data you collected, provide details about how it was collected, sampling strategy, sample size, and collection period.
3.10	Justify your data gathering decisions from 3.9.
3.11	If you needed to analyse the data before including them in the model (regardless if you collected data yourself or you used existing data), what data analysis did you do and why did you choose this specific analysis?
3.12	In what format was the data implemented? {e.g. look-up table; distribution}
3.13	Why this way?
3AQ	Any additional comments?
4	EXPERIMENTATION QUESTIONS
4.1	Describe the calibration process you followed, stating which parameters you calibrated, their ranges, your reasons, and the similarity you achieved.
4.2	Describe the experimental design process you followed, stating your reasons, and the methods you used for the different steps. {e.g. calculating warm up period, run length, and number of replications; sensitivity analysis; robustness analysis}
4.3	What type(s) of experiments did you run? {e.g. calibration; empirical validation; sensitivity analysis; performance optimisation}
4.4	For each experiment, name the purpose (objective).
4.5	Describe the parameters you used to set up the experiments?

(Continued)

4	EXPERIMENTATION QUESTIONS
4.6	Describe the data output that the model was designed to produce, your reasons for producing this output, and the data type of the output (qualitative or quantitative).
4.7	Describe the (statistical) analysis that you used on the output data and why.
4.8	Did you discuss the output with the stakeholders? What did you discuss? Why? What effect did it have on the model?
4AQ	Any additional comments?

5	EVALUATION QUESTIONS
5.1	In validation, what similarity measures did you use and why? What similarity did you get? What would you consider a good similarity and why?
5.2	How do the data outputs support an answer to the research question?
5.3	Did you discuss the validation results with the participants? What did you discuss? Why? What effect did it have on the conclusions?
5AQ	Any additional comments?

A5: Participatory Driven RAT-RS Question Suite v1.0

Filled in by (name; email): _____

Signed off on (date): _____

1	MODEL AIM AND CONTEXT QUESTIONS
1.1	If this RAT-RS use is related to a specific publication, please provide a reference to that publication.
1.2	What is the purpose of the model? {prediction, explanation, description, theoretical exploration, illustration, analogy, social interaction, or other (please specify); for an explanation of model purposes, see, Edmonds et al. (2019) http://jasss.soc.surrey.ac.uk/22/3/6.html }
1.3	What domain does the model research?
1.4	What (research) question(s) is the model addressing? {in general/in this publication}
1.5	What is the MAIN driver for your initial model development step? {theory(s), empirical evidence, existing model(s), participatory modelling data}
1.6	Explain why this MAIN driver was chosen?
1.7	What is the target system that this model reproduces? {briefly describe the target system and its boundaries}
1.8	Explain why this target system and these boundaries were chosen.
1AQ	Any additional comments?

2	CONCEPTUALISATION QUESTIONS: WHAT AND WHY?
2.1	Who did you recruit for the participatory process and why?
2.2	Describe the participatory process (e.g. environment, context, questions that were asked), including anything else that might have influenced the output (e.g. fire alarm during the participatory session).
2.3	What are the key elements of the target system?
2.4	What procedures and methods did you use to conceptualise the target system key elements? Provide a comprehensive description about how you made use of the participatory process in the conceptualization of the target system key elements. Provide details about what other sources you utilised to conceptualise target system key elements (e.g. theory, previous model(s)).
2.5	N/A.
2.6	Explain why a model element was added when this was not included in the target system
2.7	What elements of the target system were mapped into model elements? {distinguish (at least) between agents, environment, and relationships/interactions among any combination of these}
2.8	Explain why elements of the target system were included, excluded or changed in the model.
2.9	What procedures and methods did you use to conceptualise the target system elements as model elements? Provide a comprehensive description about how you made use of the participatory process in the conceptualization of the model elements. Provide details about what other sources you utilised to conceptualise model elements (e.g. use of a previous model).
2AQ	Any additional comments?

3	OPERATIONALISATION QUESTIONS: HOW AND WHY?
3.1	What data element(s) did you include for implementing each key model element in the model's scope?
3.2	Are these data elements implemented with the help of qualitative or quantitative data or further models?
3.3	Explain how data affected the way you implemented each model element and why. {i.e. explain your choice of data elements}
3.4	What are the data elements used for in the modelling process: specification, calibration, validation, other?
3.5	Why for this use and not another one?
3.6	Did required data exist?
3.7	If it existed, did you use it?
3.8	If you did not use it, why not?
3.9	For the existing data you used, provide details (a description) about data sources, sampling strategy, sample size, and collection period. For the data you collected, provide details about how it was collected, sampling strategy, sample size, and collection period.
3.10	Justify your data gathering decisions from 3.9.
3.11	If you needed to analyse the data before including them in the model (regardless if you collected data yourself or you used existing data), what data analysis did you do and why did you choose this specific analysis?
3.12	In what format was the data implemented? {e.g. look-up table; distribution}
3.13	Why this way?
3AQ	Any additional comments?

4	EXPERIMENTATION QUESTIONS
4.1	Describe the calibration process you followed, stating which parameters you calibrated, their ranges, your reasons, and the similarity you achieved.
4.2	Describe the experimental design process you followed, stating your reasons, and the methods you used for the different steps. {e.g. calculating warm up period, run length, and number of replications; sensitivity analysis; robustness analysis}
4.3	What type(s) of experiments did you run? {e.g. calibration; empirical validation; sensitivity analysis; performance optimisation}
4.4	For each experiment, name the purpose (objective).
4.5	Describe the parameters you used to set up the experiments?
4.6	Describe the data output that the model was designed to produce, your reasons for producing this output, and the data type of the output (qualitative or quantitative).
4.7	Describe the (statistical) analysis that you used on the output data and why.
4.8	Did you discuss the output with the stakeholders? What did you discuss? Why? What effect did it have on the model?
4AQ	Any additional comments?

5	EVALUATION QUESTIONS
5.1	In validation, what similarity measures did you use and why? What similarity did you get? What would you consider a good similarity and why?
5.2	How do the data outputs support an answer to the research question?
5.3	Did you discuss the validation results with the participants? What did you discuss? Why? What effect did it have on the conclusions?
5AQ	Any additional comments?