Author's response to reviews

Title: Analysing the role of complexity in explaining the fortunes of technology programmes: empirical application of the NASSS framework

Authors:

Trisha Greenhalgh (trish.greenhalgh@phc.ox.ac.uk)

Joseph Wherton (joseph.wherton@phc.ox.ac.uk)

Chrysanthi Papoutsi (chrysanthi.papoutsi@phc.ox.ac.uk)

Jenni Lynch (j.lynch5@herts.ac.uk)

Gemma Hughes (gemma.hughes@gtc.ox.ac.uk)

Christine A'Court (christine.a'court@nhs.net)

Sue Hinder (suehinderraftresearch@hotmail.co.uk)

Rob Procter (rob.procter@warwick.ac.uk)

Sara Shaw (sara.shaw@phc.ox.ac.uk)

Version: 2 **Date:** 08 Mar 2018

Author's response to reviews:

Response to reviewers on BMED-D-18-00045R1

Greenhalgh et al Analysing the role of complexity…

Reviewer #1:

I enjoyed reading this paper very much. The use of longitudinal case studies to present a concrete example of how the NASSS framework works in practice was useful.

The presentation of the framework - and its usefulness were clear to me. I admit I may be overly enthusiastic (perhaps justified) because I agree whole heartedly with the authors approach and insights around evaluation of complex eHealth programmes. But that aside - the paper is well written and the methods valid and justified throughout.

Thanks. No changes needed.

I would perhaps like even more reccomendations at the end as to how this framework could be used in practice - but there are some excellent take home messages and enough in the paper to safely use the framework in similar programmes.

We've added a stronger conclusion with firmer recommendations on page 25-27 in the revised manuscript. Note, however, that this work on recommendations is currently in its early stages and we have begun to collaborate with business and management academics with a view to testing and refining them - all of which will be addressed in a future paper.

For these reasons I would recommend accept - and I would most certainly be citing the paper myself in future work and asking my student and colleagues to read it.

I thank the authors for presenting something very complex in such a clear and straighforward way.

Reviewer #2:

I agreed to review this manuscript as it describes a conundrum that I am all too familiar with. Therefore a framework to identify the sources of complexity which besets technology implementation in health and social care services is highly attractive. The paper did not disappoint and I definitely recommend publication.

Just a couple of points for consideration;

At the outset it is stated that the data used were derived from two large scale programmes that the authors were engaged in. However a number of the case studies did not appear to be associated with these programmes. The acknowledgements at the end suggest that all but one were. Elsewhere there is reference to other colleagues applying the framework to their work. This needs clarification - how was each case identified and from where? I presume that the framework used secondary analysis of data from each study and did not involve any additional data collection - I think that it is important to be clear about this for those of us who may wish to experiment with the NASSS framework.

Actually the paper is reporting a primary, prospective study and all the data relates to the case studies that were in the original sample. We've clarified this on page 9 (thanks for pointing it out).

I realize that it is early days but within the discussion/conclusion I do think that there is room for two or three sentences on how this might be taken forward by commissioners in particular.

Reviewer 1 made this same point, so we've noted it carefully and done some more work! See new recommendations on page 25-27.

Reviewer #3:

An excellent and comprehensive mountain of work, synthesised and accessible for academic, clinical and management readers.

I have two minor concerns with respect to the readers' take home messages;

- in reading the narrative summaries of each domain in the results, could more emphasis be included in distinguishing between complicated and complex scenarios - given this seems to be the key predictor of outcomes, can you highlight more specifically those examples that are complicated vs complex? Is there a way you could visually tabulate this?

Good point, we have now done this. See in particular pages 13-21 in the revised document.

- can you include a summary take home message at the end of the Discussion, to indicate whether stakeholders can influence projects positively by focusing on and addressing complicated components to foster their chance of success? I might be too pragmatic in my search for the so-what of this article - but it does appear that in some 'complex' situations, they can be addressed in part by identifying and systematically working through levels of complicated design. Do you have data to suggest this? If so, what are the next research steps in this space?

See above. We've taken the Discussion/Conclusion forward on page 25-27 with further recommendations and 'next steps', notably in relation to applying principles for supporting non-linear change and understanding, reducing and 'running with' complexity.

Reviewer #4:

The authors apply their NASS framework to six cases of technology programs in this paper. I learned and thought a lot upon reading it. I do however have some thoughts. I apologize in advance for any of my own misunderstandings of NASS itself or complexity theory, as I am not an expert in implementation or complexity theory per se. Happy to be corrected if needed. My comments are as follows.

This commentary was particularly helpful as it seems that the first three reviewers are experts on complexity whereas reviewer 4 admits he is not, so his comments are crucial in helping us to revise the paper for a less expert audience (which is the whole point of the BMC complexity series). So thanks for being honest and also pushing us to do a bit more work.

Intro/overall

1. First, I have questions about the overall premise of the activities described in the paper. The following may seem abrupt, but is really just a request for clarification.

a. First, what does NASS framework do, or rather why does it exist? The process of creating it is noted in the introduction of the paper, but what is not outlined is the intended purpose of the NASS framework. What broader questions is it intended to answer? The component questions for NASS listed in the paper are mainly descriptive; my question is about what we are to do with the NASS framework other than describe, or, alternatively, what answering those descriptive questions is supposed to allow us to do. Not saying it shouldn't be used, but rather looking for a clear justification. Considering that there are other theories and frameworks related to patient burden or patient experience, provider adoption, implementation, etc. (some of which informed the NASS framework, but one could also think about RE-AIM, extensions of the models cited and other literature), the reader needs some clarification as to what gap the NASS framework was intended to fill. This, to me, would be more useful than the current description of its creation process.

We accept this criticism and have included a description of what the NASSS adds over and above other theories and frameworks (page 4), and preliminary suggestions for how the NASSS framework might be used (page 25-27). (We did write an earlier theoretical paper [1] which I think the other reviewers may have been familiar with, but we can see that this paper needs to be a bit more freestanding).

b. Following that, the reasoning should also be clarified for why the authors then used the NASS framework, and connections to complexity theory concepts, to organize descriptions of the different technology programs which themselves were used to develop NASS itself. Beyond a demonstration exercise/illustration (which is not wholly uninteresting), I would be more motivated in my interest if the authors could also show: what does this doing this enable us to learn in a substantive sense? As a reader, I want to know why I'm going along for the ride. Will this paper get me toward a deeper understanding of the common mechanisms or processes whereby complicated or complex arrangements drive de-adoption or failure in various forms, or just a closer understanding of the NASS framework itself? I would think both, and stating so would whet my appetite more successfully.

We hope the extensively revised paper now makes it clearer why the reader is being invited along for the ride.

c. In relation to the paper's premise: Why should I assume that the simplicity-complexity spectrum/taxonomy is the driving feature behind failures of technological innovation rather than easier-to-study constructs like patient experience? Again, all of this gets at the need to see why the illustration described in the paper is useful.

The short answer is the complexity came from our data. We didn't begin with complexity as the dominant (or even a minor) component of our explanatory models. Indeed, when we first started studying patient-facing assistive technologies, we focused almost exclusively on the patient experience. Our paper 'What matters to patients with assisted living needs…' used a philosophical lens from Heidegger and phenomenology to theories and study the patient experience.[2] That study reflected (and perhaps extended) dominant thinking in north America - and indeed the paper won the annual prize from the American Medical Informatics Association! We have also published MANY other studies of the individual patient and staff experience of technology adoption and use (see for example [3-10]). However, on completion of these various studies we felt we'd missed out a major theme in our own data - the ORGANIZATIONAL challenges associated with introducing and supporting assistive technologies. The SCALS program, funded by a major award from Welcome Trust, explicitly shifted our unit of analysis from the individual patient/client to the organization.[11] Furthermore, whilst we put organizational challenges on the SCALS grant application, as the case studies unfolded we realized that commercial / regulatory issues were looming large in some case studies and had to be factored in.[12] And another CRUCIAL finding in our data (not identified by other authors previously) was the unpredictability of many diseases and conditions, to the extent that a high proportion of patients were being classified as "unsuitable" for an algorithmic technology. It was relatively late in the day that the idea emerged to use simple/ complicated/ complex to get a handle on these and other dimensions of the case studies, and

when we did that, the data all fell into place (mostly!). Sorry for the long explanation, we've inserted a single sentence to reflect this at the top of page 6.

2. Opening statements in the introduction set up a narrative that seems to imply that problems with technological innovation are unforeseen. A number of theories or models, some cited by the authors, would tell us that's not the case. Policymakers may indeed think it is great, but setting up the narrative in this way seems to imply that the NASS framework is the only game in town. I would rather see why it's the most useful game in town by saying "other frameworks have done X, but NASS does Y and that's important because Z." So, this is related to point 1a above, but is specifically about the narrative laid out from sentences 1 and 2.

This was in the theoretical paper on NASSS, published previously. We've included a brief explanation on page 4. In short, we - with acknowledgement - wove together selected previous frameworks and added additional domains.

3. The last two sentences of the introduction could use a little rephrasing in my opinion. Maybe I am being thick-headed, but words like "secondary research" and "empirical findings," used seemingly in contrast, don't help me understand the difference between the two papers. Another way to do it (which might also help with my earlier points in 1b) could be to explicitly state at the end of this section what the research aim or research question of the current paper is.

Good point. Done (page 8).

Complexity theory section

4. The complexity section is good. However, the outline of what might be "complex" in some of the domains, at least based on my muddled understanding of complex adaptive systems, seems like it could use more thought. For instance, a complex illness fits what I have in my head regarding complexity-it is explained to me clearly what would make an illness "complex." However, a number of the other domains list the dimensions of the domain (e.g., material properties of a technology) rather than what would actually make them complex. Having hard-to-understand material properties does not seem to fit the idea of complexity, so what would? Having unpredictable material properties made of multiple elements that bounce off of each other in emergent ways? What would that look like concretely? Similarly, having significant barriers doesn't mean something is necessarily complex in domain 7 if said barriers operate pretty mechanically, or simply. It would take more to make them complex,

wouldn't it? Again, this occurs in a few domains, so just outlining a little better what complexity would look like would help me wrap my head around the marriage/synthesis of complexity theory with the NASS domains. Maybe it is the case that some domains (e.g., organizational setting) are easier to apply complexity theory to than others. If so, clarifying that might help my understanding as well.

This was a really helpful comment. We've sharpened what we mean by 'complex' (pages 6-8).

Context section

No real issues here. My points 1b/1c and 3 above, though, remain after reading this section. If these data sources informed the development of the NASS framework, would it not have been more useful in this paper to "validate" the model by applying it to other cases while also explicating its use? If better justified (in response to 1b/1c and 3), this answer might be moot, but this decision to use the same data sources is a methodological decision, so some justification as to its methodological usefulness would help.

Yes, sort of, but "validating a model" is the sort of language we use when the model seeks to be predictive and implies a Newtonian universe. Instead of that, we've indicated in various sections in the paper that the work is ongoing and the NASSS framework is continually being improved (rather like Carl May is doing with normalization process theory - he treats it as a 'work in progress' to the extent that he holds a conference every couple of years to see how people have applied and adapted it).

Case example descriptions

No issues here.

Analysis section

6. The details on this analysis section are a bit thin.

a. The first sentence suggests that these were a theoretically sampled set. Does that mean there was some larger set out there from which these were sampled based on the NASS domains? If not, are these not just a convenience sample of cases which were then arrayed according to their usefulness in explaining domains?

Not quite. We began with two case studies in which the practitioners/policymakers had approaches us. We added two more which were 'on our doorstep' (obvious things to study). We then added two more to provide contrast (for example, all the earlier ones were patient-facing but we felt we should look at least one 'back office' technology). Our steering group encouraged us to develop what they called a "typology" of the technologies and cases. NASSS was the product of that work. In sum, not quite a convenience sample but certainly aided and abetted by convenience. We've made this clear on page 9.

b. I could use a little more detail as to how the qualitative data were analyzed thematically. How broad were such themes to start with, were they relatively inductive, organized a priori around domains, etc.? What kinds of codes, notes, memos, or other approaches were used to organize themes, or to guide communication around themes among team members?

We were doing a systematic review in parallel, so we didn't approach the data 'blind'. Rather, we assessed the extent to which our findings mapped to existing frameworks. When they didn't, we developed additional domains. The obvious new domain was Domain 1 (the condition or illness) which was VERY prominent in our data but had not featured at all in previous frameworks (except an earlier one developed by our own team [8]). This was explained on page 13.

c. It seems like team meetings were integral. What was the timing of said meetings in relation to, for instance, coding/analysis activities, the systematic review, and other pieces of this process?

Formally, we met monthly (we've added that on page 13). Informally, we are a close-knit research team working mostly from shared offices. We talk about data all the time. Steering group meetings were six-monthly and provided us with a useful external steer. We don't think the timing of meetings is relevant to the paper - indeed I'm a bit worried that this push to 'rationalise' the work we did is going to create a distortion of the messy way in which it actually unfolded. Far more important than who met and how frequently or how a particular code was applied was the wider culture of our team and the interpretive and reflexive approach we took to developing the case narratives. You could say we used the 'hermeneutic circle', developing a provisional case narrative early in the study and adding new data iteratively to enrich the narrative.

d. How exactly were quantitative data woven in, and at what stage and form?

See response to previous question. There's nothing "exact" (in the technical sense) about how this was done. There are very often numbers in stories, which are simply part of the narrative (e.g. The grand old Duke of York who had ten thousand men). Numbers get inserted into the case narrative when they help the story make sense. See my paper on this topic, 'Stories or numbers or stories and numbers' (https://www.kingsfund.org.uk/sites/default/files/media/GREENHALGH%20Stories%20or%20numbers%20or%20stories%20and%20numbers.pdf).

e. What is meant by "data trends"?

Trends in data over time. So for example, looking at whether the uptake of video consultations by patients was increasing or decreasing over time. I'm not clear why the reviewer finds this unclear (but happy to be guided further).

Results

7.

a. Complexity of condition certainly seems relevant. However, to my untrained mind, it might be useful to discuss whether/how one can separate out aspects of what makes something complex versus other constructs which can be broken down more easily, such as degree of risk, comorbidity, and history. Yes, all these can come together in a complex fashion, but I don't know that listing them gives me a good sense of what truly makes someone a complex case. Are we to understand that having condition a itself (e.g., heart failure) will make someone complex, or is more required?

We've added in a clearer explanation of what makes something complex rather than just complicated (pages 6-9). But we need to be clear, complexity of a phenomenon is a) to some extent a subjective judgement and b) an emergent phenomenon, hence something might be complicated and then become complex. Again, we can't apply technical methods to this interpretive challenge. Nevertheless the reviewer is right to push us to get the broad descriptors tighter so people can make more consistent judgements.

b. A similar pattern (e.g., high/low risk used as complexity feature) can be seen in some other domains, such as technology, value proposition, and others. I'm wondering if the authors could translate risk from being high/low in various forms into something that would more clearly characterize (what I understand at least) about complexity: uncertainty. In some cases, if high or low risk is "known" by actors, they might react to it in a relatively simple (or at

least complicated) manner. However, if they are reacting to the uncertainty itself, that would seem to be more toward the complexity side. Maybe I'm off base, but at the very least, some talk either in findings or discussion about high risk, uncertainty, and complexity would be a useful takeaway. That's to say nothing of severity of an outcome-e.g., something might be high risk for a low-impact event. Anyway, a teasing out of those concepts would be fun to read, and useful.

Again, this was a good steer because it helped us clarify in each domain what we mean by complex (see page 6-9).

c. Workloads and barriers seem relevant to complexity. But, like risk, I'm wondering exactly how this makes something more complex. A barrier may be responded to simply/mechanically, couldn't it? Is it that barriers and the tasks involved in workloads create more moving parts? Again, explicating the mechanism by which the various features of domains drive complexity versus simplicity would be useful, partly explicitly here in the results probably and then with some call-back in the discussion.

The barrier metaphor is, we agree, a red herring and we've removed it. Let's take what we described in the initial manuscript as the "biggest barrier" to IT uptake in healthcare: staff resistance. The reason that barrier can't simply be "lifted" is that staff resistance is itself a highly complex and often deeply institutionalized construct. Staff resist not because they are obtuse or lazy but because - for example - the technology is (they feel) forcing them to be unprofessional (e.g. by sharing patient data). We've clarified this point on page 17. (Note that organisational issues are NOT mainly about workloads; they're about more complex things such as interacting routines as we explain in the paper).

d. The one thing that I felt I didn't see enough of was the way in which complexity in different domains interacted; that would have been perhaps the fullest expression of complexity under the NASS framework and the closest expression of how things happen in regular practice. I didn't really see it in the domain descriptions in the results, or in a separate results subheading, or really in the discussion. I think these authors could certainly speak to it.

This is what the narratives illustrate, surely. We've emphasised that point (page 23).

Discussion

5. The opening part of the discussion does more for me than did the background or results in actually showing how domains connect to complexity. That's probably by design. However,

the fact that I am set up from the beginning to look for complexity in the domains means that I am left wanting more of this kind of outline up front. Perhaps explain early on that you're going to work through domains first and then get into complexity more in the discussion, or find more ways to make clear (e.g., in response to points 7a-c) in the results how the dimensions/features of programs on those domains (e.g., high risk) make something complex. There were some exceptions. For instance, mentioning interoperability of systems in results for technology, and then calling back to it in the discussion, helped me see that point. But I felt that the discussion generally spelled things out in a way that was more understandable for me than did the results.

Great feedback thanks. We've added more on complexity on pages 6-9 so the reader is less at sea. We already did say in the first sentence of the Results section that we were going to work through NASSS and then talk about complexity in the Discussion. We've clarified this sentence further (page 13).

9. The two statements after the domain descriptions (p 22, lines 9-11 and 15-23) seemed less well-supported by the data/analysis than some other statements in the discussion.

a. lines 9-11 (Complexity on multiple domains, outcomes less predictable, etc.): I don't know that this kind of spectrum/dose-response effect is clearly identified in the data/results, is it? If so, I missed it. It's a very quantitative statement to make. Seems a truism. Is it useful?

It's an interpretation. We're not making a statistical prediction. "Dose-response" is from a very different repertoire. We've tweaked the sentence (page 23) to make it clearer that we're not implying a quantitative or predictive conclusion.

b. lines 15-23 (planners planning for complicated, not complex). This may be in your data, but I don't remember seeing anything results indicating that this assumption exists across cases. Maybe this was because findings were organized by domain. Or, maybe the authors are using discussion of this and examples as they follow in lower lines to extend the paper-if so, I might suggest some broader, cross-cutting findings listed in the results section itself, such as this and perhaps some of the supporting evidence for complexity in multiple domains resulting in lower amenability to planning/implementation, if available.

I'm a bit concerned that this reviewer is seeking to deconstruct our findings in a way that isn't helpful or meaningful. We can't say "finding A in case B plus finding C in case D led us to higher-order finding E". To draw such a linear link would negate the whole genre of the study which is to surface and explore the messiness of multiple interacting influences. It's a defining

feature of complex systems that they don't behave logically. We hope we've made this clearer throughout the paper in this revision.

c. A similar issue with p. 23, line 15-29 (statement about users experimenting). This is super interesting as a finding but is organized with the discussion.

That comment in the discussion is picking up findings that were explained in some detail in the Results section (page 21-22 from "The material features…" through to "…fit for purpose").

Trish Greenhalgh on behalf of all authors

3rd March 2018

1. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond Adoption: A New Framework for Theorizing and Evaluating Nonadoption, Abandonment, and Challenges to the Scale-Up, Spread, and Sustainability of Health and Care Technologies. J Med Internet Res. 2017 Nov 01;19(11):e367. PMID: 29092808. doi: 10.2196/jmir.8775.

2. Greenhalgh T, Wherton J, Sugarhood P, Hinder S, Procter R, Stones R. What matters to older people with assisted living needs? A phenomenological analysis of the use and non-use of telehealth and telecare. Social science & medicine (1982). 2013 Sep;93:86-94. PMID: 23906125. doi: 10.1016/j.socscimed.2013.05.036.

3. Swinglehurst D, Greenhalgh T, Roberts C. Computer templates in chronic disease management: ethnographic case study in general practice. BMJ Open. 2012;2(6). PMID: 23192245. doi: 10.1136/bmjopen-2012-001754.

4. Wherton J, Sugarhood P, Procter R, Rouncefield M, Dewsbury G, Hinder S, et al. Designing assisted living technologies 'in the wild': preliminary experiences with cultural probe methodology. BMC Med Res Methodol. 2012;12:188. PMID: 23256612. doi: 10.1186/1471-2288-12-188.

5. Procter R, Greenhalgh T, Wherton J, Sugarhood P, Rouncefield M, Dewsbury G. The ATHENE Project: The importance of bricolage in personalising assisted living technologies. International Journal of Integrated Care. 2013 (7).

6. Greenhalgh T, Stones R, Swinglehurst D. Choose and Book: a sociological analysis of 'resistance' to an expert system. Social Science & Medicine. 2014;104:210-9.

7. Procter R, Greenhalgh T, Wherton J, Sugarhood P, Rouncefield M, Hinder S. The Day-to-Day Co-Production of Ageing in Place. Computer supported cooperative work : CSCW : an international journal. 2014;23(3):245-67. PMID: 26321795. doi: 10.1007/s10606-014-9202-5.

8. Greenhalgh T, Procter R, Wherton J, Sugarhood P, Hinder S, Rouncefield M. What is quality in assisted living technology? The ARCHIE framework for effective telehealth and telecare services. BMC Med. 2015;13:91. PMID: 25902803. doi: 10.1186/s12916-015-0279-6.

9. Wherton J, Sugarhood P, Procter R, Greenhalgh T. Designing technologies for social connection with older people. In: Prendergast DK, Garattini C, editors. Aging and the Digital Life Course Oxford: Berghahn Books; 2015. p. 107-24.

10. Greenhalgh T, Swinglehurst D. Studying technology use as social practice: the untapped potential of ethnography. BMC Med. 2011;9:45. PMID: 21521535. doi: 10.1186/1741-7015-9-45.

11. Greenhalgh T, Shaw S, Wherton J, Hughes G, Lynch J, A'Court C, et al. SCALS: a fourth-generation study of assisted living technologies in their organisational, social, political and policy context. BMJ Open. 2016;6(2):e010208. PMID: 26880671. doi: 10.1136/bmjopen-2015-010208.

12. Greenhalgh T, Shaw S, Fahy N. The bright elusive butterfly of value in health technology development. International Journal of Health Policy and Management. 2017;6(x):1-5.