SHORT TERM SCIENTIFIC MISSION (STSM) – SCIENTIFIC REPORT

STSM participant: Marta Natalia Wróblewska
Home institution: Centre for Applied Linguistics, Warwick University (UK)
Action number and name: CA15137 European Network for Research Evaluation in the Social Sciences and the Humanities (ENRESSH)
STSM title: Delineating the field of SSH impact: lessons from an expert survey STSM
Start and end date: 12/07/2018 – 01/08/2018
Host institution: Centre for Higher Education Policy Studies (CHEPS) University of Twente (Netherlands)

Aim OF THE STSM

The purpose of the STSM was to contribute to establishing a common vocabulary and understanding of what constitutes research societal impact. This task is important in view of the fact that the evaluation of societal impact of funded research has become a crucial and controversial topic both in the academic debate and in policy-making. In the framework of the STSM a previously-conducted expert questionnaire was to be analysed by the grant holder in order to provide an analysis of the most common definitions of key concepts within SSH valorisation. Additionally, new research questions, regarding the construction of objects of knowledge and expertise emerged during the analysis and were addressed.

The STSM had also a function in advancing the career goals of the applicant, an early career researcher (ECR), by offering opportunities of networking, exchange with colleagues from a neighbouring discipline (science policy studies), presentation of own previous findings on the topic to expert audience, and – crucially – supervision from a senior colleague, which allowed the ECR to develop her skills in policy data analysis and academic writing and allowed for a stimulating exchange on career opportunities in academia.

DESCRIPTION OF WORK CARRIED OUT DURING THE STSM

Work carried out during the STSM proceeded along the lines indicated in the work plan included in the application for the grant:

Week 1-2 – analysis of surveys collected by ENRESSH team, together with PI Paul Benneworth. Cleaning and preparing the data, coding using MAX QDA software, analysing codes, selecting themes to be explored in depth, theory-building

Week 3 – drafting report on findings (to be subsequently written up as a CHEPS working paper), planning publication based on findings

Additionally, on 24.07.18, the grantee gave a lunchtime seminar to colleagues at CHEPS, presenting her existing (PhD) work on the topic of impact evaluation. She received valuable feedback, particularly on the directions of future, comparative studies which might include the Dutch context.

DESCRIPTION OF THE MAIN RESULTS OBTAINED

1.Background

Recent years have brought an increasing pressure on impact evaluation as goal for science policy and evaluation systems. Although protocols for impact evaluation are being developed in several countries there is no agreement as to the definition and operationalisation of 'impact' (Donovan, 2017). Hence there is a process of constant re-defining 'impact' depending on the context in which it is to be used (e.g. national, disciplinary, STEM vs SSH, ex ante vs ex post evaluation etc.) As a consequence, we can observe two contrary tendencies: on the one hand a desire to define impact "once and for all" and to establish the 'best' way to evaluate it, and on the other hand – a tendency towards opening the existing approaches up for debate, broadening the definition, translating the notion into local contexts etc. On the crossroads of these conflicting tendencies, interdisciplinary and international communities form around the task of addressing the pressing issue of 'impact evaluation'. Such communities – ENRESSH being one of them – can be considered as meaning-makers and knowledge-generators, but their members do not only study and describe a field – they actively construct, delimit and define it.

2.Research question(s)

While the initial aim of the STSM and the data analysis it involved was to further the understanding of existing concepts and definitions of 'impact', in course of the investigation the researchers noticed that alongside processes of defining and classifying impact, there are also processes of *establishing expertise and legitimacy* which can be traced in the data. We therefore decided to broaden the scope of our investigation to analyse the questionnaire not just from the point of view of the discursive construction of a particular object of knowledge ('impact') but also of the parallel construction of expert positions on the topic. Hence, following the preliminary overview of the data, we approached the analysis with two connected research questions: 1) how, on the basis of what existing points of references and repertoires is 'impact' constructed? 2) how do those who define it construct their own position as experts in the process?

The STSM involved a qualitative analysis of data from an open-ended questionnaire, which had been created and circulated in 2016. The questionnaire was developed using a co-creation methodology with a selected group of experts within the COST WG network and then refined by circulating it to get comments from all participants. The final questionnaire included a total of 13 open questions each of which contained a number of prompts, as well as a final open question allowing further remarks, comments on the questionnaire itself as well as anything that the respondent felt necessary. A total of 29 questionnaires were returned, from 15 countries and from a broad range of fields within SSH, as well as from stakeholders who came from non-research positions within universities, science councils and other participants. These 29 returns form the basis for the analysis carried out during the STSM.

4. Analysis and findings

4.1. Opening the black box of impact

We looked at the questionnaire data as the surface of discourse produced by an epistemic community in the making (or a pre-epistemic community). In the absence of cut and dried theory (Donovan 2017) we are observing theory in the making. We start the analysis with an observation that there are two complimentary and parallel processes at hand : 1) construction

of the object of knowledge (and definitions around it), 2) construction of expertise and "legitimacy to know and to make claims" (both of the responding individual and of the ENRESSH network). We looked at the emerging keywords, definitions, points of reference, discursive repertoires to understand where certainty is emerging. We examined processes of expertise construction to understand the role of context and contingency in the process of producing knowledge and expertise – see diagram 1 below for visual representation of the described processes, together with notions which will be used in the analysis which follows.

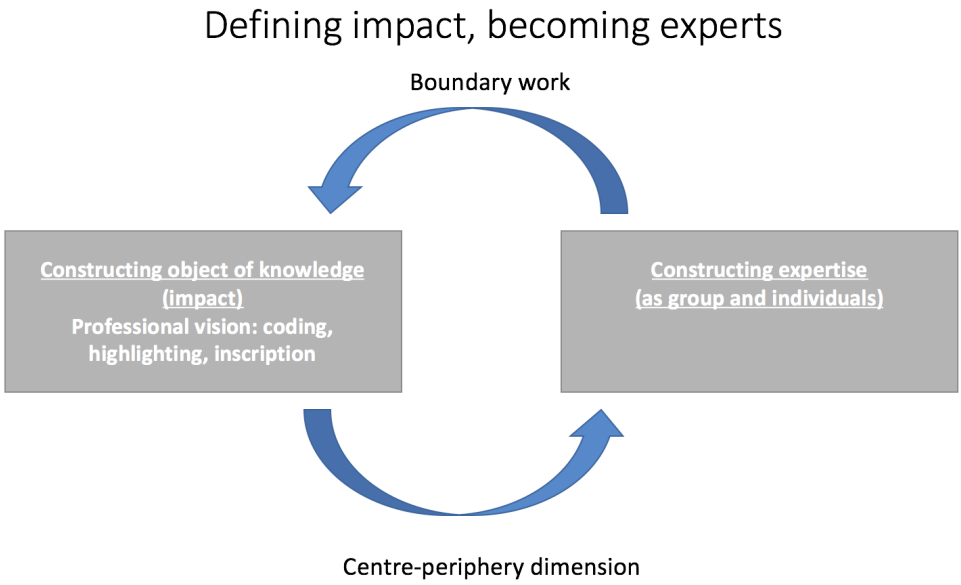

Diagram 1 – Defining impact, becoming experts

Rather than offering final definitions of impact or solutions to the problem of its evaluation, in this study we are interested in opening the 'black box' of knowledge construction to show how knowledge claims emerge out of epistemic communities which have their own dynamics, tensions and power struggles (rather than being neutral 'knowledge factories'). This observation, and its exemplification in the data, may be of consequences for directions of thinking on impact evaluation, and research evaluation more broadly.

The following paragraphs explain the two observed dimensions of discursive construction in more detail.

4.2 Defining impact

What comes to the fore when analysing the use of the term 'impact' in our dataset is its multidimensional, 'archipelagic' aspect. While WG2 of ENRESSH (the respondents of the questionnaire) is a community very much built around the notion of 'impact', the respondents were far from agreeing on just one definition or one understanding of the concept, let alone pointing to a single, 'ultimate' way of evaluating it. Rather, members of the community, while taking part in a broad debate (of which the questionnaires can be understood as a 'surface'), point to several core notions, keywords, firm points of reference but also more 'fuzzy' areas of interest and importance. Ultimately, this process leads to identifying the pieces of evidence, points of reference and repertoires which will become salient, significant (or symbolically influential, 'totemic') for the community. We have identified several points

of reference or repertoires which were tapped into repeatedly by different respondents: existing research on impact, public debates, own studies on subject, existing protocols and approaches, personal experiences (often narrated as stories or anecdotes). Two vectors can be spotted here – an 'inwards' one which tends eventually towards circumscribing and defining impact in an operationable form (this tendency is visible in the respondents' pointing to the existing commonalities, attempting to offer definitions, proposing cut and dried theories), and an 'outwards' tendency which aims at broadening the concept and the debate around it (visible in the highlighting the disagreements, dangers, challenges, areas of uncertainty, issues which are up for debate).

We conceptualise the process of meaning-making around the concept of impact using Goodwin's pragmatic notion of 'professional vision'. Goodwin writes: "discursive practices are used by members of a profession to shape events in the domains subject to their professional scrutiny. The shaping process creates the objects of knowledge that become the insignia of a profession's craft" (Goodwin, 1994). We argue that such a process of shaping 'events in the domain subject to (our) scrutiny' is at hand in the analysed data. The process, as argued by Goodwin, proceeds in three steps: 1) highlighting, 2) coding, 3) inscribing. The first step relates to subjects' highlighting of the broad area of events, behaviours and other phenomena which fall under impact (this is done by listing keywords, telling stories and anecdotes, bringing up possible areas where impact can occur). 'Coding' occurs when respondents refer to particular definitions or make categorical statements about what is and what is not impact – an action which may be analysed using the notion of 'boundary work' which has been often used in Science and Technology Studies (Gieryn, 1983; Star Leigh, 2010; Star Leigh & Griesemer, 1989). Finally, the third step – inscription – consists in producing visual representations, schemata, diagrams, graphs but also creating/pointing to technologies which may help 'house' or 'categorise' (instances of) impact.

4.3. Centre-periphery issue

Respondents of the questionnaire almost universally show high levels of knowledge on the problems related to impact evaluation (visible in awareness of the recent debates, existing approaches, their respective challenges and a deep reflection about the consequences of introducing systems of impact development). However, we notice an emergent centre-periphery dynamic, which can be linked to the phenomenon of "two-speed impact development". While certain countries or institutions have already developed quite sophisticated systems of impact evaluation and are currently at the stage of refining and fine-tuning these methods and protocols, other countries and institutions are still gauging the ground for possible impact evaluations and debating their future shape. Therefore, there emerges a 'hard core' of most recognizable, most used and most copied solutions, with the systems of REF (UK) and SEP (Netherlands) at the centre, and other proposed protocols and approaches (Siampi, Pathways to impact, Hera value etc.) surrounding the centre. The contrast between the relative stability of the established systems and the 'fuzziness' of those which are only now being developed – often somehow on the basis of SEP and REF (this is the case for French CENRS approach, University of Ghent's protocol, the Norwegian and Swedish impact evaluations) – give rise to flow of expertise based on a centre-periphery dynamics. The recognition that different countries and institutions are at different stages of development when it comes to impact evaluation leads to a focalization of discourse around the existing solutions which, often without the explicit will of the participants of the debate, become fashioned as 'models' to be copied and imitated. This polarization of discourse along a centre-periphery line also affects the position from which the different locutors in the

discourse speak. This brings us to the issue of power dynamics in the emergent expert community, an issue which has been captured in a poignant manner by one of the respondents who argued "Drawing rules around what is impact and what is not will wind up alienating some academics".

4.5. Constructing expertise

Considering the high levels of uncertainty which are endemic to any pre-epistemic community, definitions can be contested and unpicked but ultimately have to be collectively adjudged as being legitimate so community participants seek to signal in various ways, the legitimacy of what they are arguing as having a wider salience or validity for the group as a whole. So, the second element of the model we propose focuses on the dynamics by which the participants legitimate their definitions of impact – while co-constructing the object of knowledge they also construct their positions as experts.

Similarly to the two vectors mentioned in section 4.2 which related to defining, narrowing down the notion of impact on the one hand and opening it up for debate on the other hand, also here we can spot two opposite tendencies. Some responses seem to signal lack of certainty, impossibility of taking a stance (linguistically signalled by hedging devices – "I think; to me; does it makes sense?) when other responses, on the contrary, suggest high levels of certainty (linguistically signalled by the using prescriptive verbs – *we must, we need to* – boosters, but also by the more academic register of the response, and the use of academic references, particularly to one's own work). Hence, the respondents' knowledge of practices, awareness of debates, awareness of peer-reviewed publications, authorship of peer-reviewed publications, having undertaken research as impact serve not only as points of reference for the construction of a new object of knowledge ('impact'), as described in section 4.1. but also as building blocks of one's own authority claims.

Also group expertise and group legitimacy (of ENRESSH, an emergent epistemic community) is discursively constructed in the process of debating the definitions of impact and protocols for impact evaluation. This process can be traced in the data for instance in references to 'tacit knowledge' – acronyms (TTO), theories and approaches ("Callon's techno-economic networks"), established authorities (impact a la Bozerman) names of entities and projects Frascati) and groups ("the Edinburgh group") as well as personal acquaintances ("this is a comment that [name] made me years ago"). The construction of a group identity can be also seen in the strategic use of pronoun "we" often paired with the prescriptive verb "must / have to". But the use of "we" in the answers to the questionnaire is far from uniform and does not always refer to the ENRESSH network – respondents use it to talk of their national context ("we in France"), their academic affiliation ("we at the university of Ghent") and membership in other (often impact-related) projects ("we in Siampi") – this shows how a new group of experts is being actively constructed on the crossroads of existing and often overlapping groups of interest, while different identities are played out to forge new and convincing expert positions.

Hence, in the process of constructing, defining and 'naturalizing' the concept of impact also certain positions on the issue emerge. The ones which are held to be legitimate and compelling will be accepted and incorporated into the sense of what matters whilst those that are not are discounted. In this process participants who are able to legitimate, compel and convince likewise see themselves become more central actors in the group whilst others may

disengage or detach if their ideas are not taken forward. We have observed the consequences of this process in the fluctuations of the ENRESSH network membership – for instance one of the original respondents of the questionnaire, in email communication, explained that they reduced their engagement with the group, as they did not agree with the 'mainstream' attitudes, standpoints and values which were coming to the fore in the group's communication (again suggesting a crucial importance of the centre-periphery dimension of knowledge construction).

It might be the case that every community, as it becomes more consolidated, particularly around certain shared concepts and knowledge also draws its own boundaries, leaving some subjects outside. However, with this study, we are trying to make this process transparent and traceable in order to reflect on it and address the possible bias which might ensue.

4.6. Conclusions

The 'area' of impact is targeted by governments, policy-makers and – to a degree as a consequence – by groups of academics and practitioners in the field, contributing to a problematization of the practices which can be grouped under the label of 'impact'. Despite the efforts of groups of researchers – like the ENRESSH network members and the various research projects which proceeded it, or work in parallel with it – the notion remains fuzzy, difficult to pin down and subject to often opposite tendencies: one of defining and constraining the concept and the other of constantly questioning and broadening it, as it becomes applied in new contexts. This observation on the generation of knowledge and certainty around the notion of impact as well as the practice of impact evaluation can be interpreted in light of Foucault's reflection that "rules are empty in themselves, violent and unfinalised, they are impersonal and can be bent to any purpose" (Foucault, 1984, pp. 83-86).

4.7. Policy findings:

While on the side of policy-makers there seems to be a request for definitive and standardized definitions, metrics, indicators and protocols around research impact evaluation, we argue that these can never be as 'solid' and 'robust' as policy-makers would like, as they are always the effect of discursive co-construction. This observation is not meant as a critique of existing impact evaluation protocols nor an argument for abandoning attempts of developing systems for impact evaluation, but as a reminder of the contingent nature of the concept and of the practices which emerge around it.

Hence we stress the importance of having a transparent and open debate in every context where impact is to be introduced, rather than somewhat mechanically 'importing' existing solutions from foreign contexts (Wróblewska, 2017a). Indeed, we consider the quest for a single, comprehensive definition of impact a waste of time. In contrast, the time invested in holding a broad debate around impact will allow the notion to be constructed in harmony with the local ontologies (tacit assumptions about the role of research in society, the nature of scientific work etc.). It will also allow for a reflexive embedding of the notion in the communities which will use it.

The purpose of 'opening the black box' is not merely to carry out a self-reflexive (somewhat self-indulgent?) exercise within the community, but above all to argue that impact is not a

'objective reality' (as often presented in neutralizing, reifying policy documents), but a concept constructed by epistemic communities, which draw on different, local repertoires. Therefore, we argue, that policy-makers should not simply draw on existing, tested solutions as these will likely be at odds with the specific academic values of the local context (Wróblewska, 2017a, 2017b). Therefore, a debate is always needed, even if only to redefine, localize the same basic concepts, building a 'local ontology of impact' (mezzo-ontology). Our study shows the contingency of this process and the lack of convergence to a single definition.

In highlighting the need for a transparent debate, we do not point just to the necessity for observers (academic community) to follow the debate around the notion of impact and to trace its final conclusions, but also the urgency of self-checking and self-reflection of the 'experts' who do engage in the debate. In the 'paradigmatic' systems of impact evaluation (Norway, Netherlands, UK – NONLUK) such reflexivity would involve constantly reminding oneself that impact is a construct and must be regarded critically, while in the non-paradigmatic impact countries (NPIC) it would be a question of not following unreflexively the example of the central systems. In the core of the emerging impact community in turn (ENRESSH and cognate groups) reflexivity would involve a constant consideration of the ongoing tension between centre and periphery, which affects the discourse on impact and the positions of experts. We must remember that even in relatively non-hierarchical networks power struggles over legitimacy take place, and the centre-periphery distance / tension remains present. The only way to avoid it is to continue reflexively opening up the dialogue.

FUTURE COLLABORATIONS

The STSM holder and the PI plan to continue collaborating on preparing publications drawing on the findings from the expert questionnaire and, more broadly, in the framework of the ENRESSH Network.

PLANNED PRESENTATIONS AND PUBLICATIONS

On the basis of the findings from the research carried out in the framework of the STSM the PI and the grant holder intend to prepare two publications:

1. CHEPS working paper (preliminary date of submission to editors – 15.09.18)
2. Journal article – to be submitted to suitable outlet (Science as Culture, Minerva, Research Evaluation…) (preliminary date of submission– 15.01.19)

Publishing first an open-access, online working paper will allow us to recieve feedback from the community which we are describing. This is important both for ethical reasons (maintaining an ongoing relationship with the studied respondents and allowing them to comment on the analysis we are advancing) and for methodological ones (to make sure the analysis we are advancing reflects the reality of knowledge-building in the network as perceived by its participants).

ANNEXES

1. Map of concepts and coding (mind map) generated from the analysed data produced using MAX Qda software (png and pdf format)
2. Codebook used for coding questionnaire data

References

Collins, H. M., & Evans, R. (2002). The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science, 32*(2), 235-296. doi:10.1177/0306312702032002003

Donovan, C. (2017). For ethical 'impactology'. *Journal of Responsible Innovation*. doi:10.1080/23299460.2017.1300756

Foucault, M. (1984). Nietzsche, Genealogy, History. In P. Rabinow (Ed.), *The Foucault Reader* (pp. 76-100). New York: Pantheon Books.

Gieryn, F. T. (1983). Boundary work and the demarcation of science from non-science: strains and interests in professional ideologies of scientists. *American Sociological Review, 48*(6), 781-795.

Goodwin, C. (1994). Professional Vision. *American Anthropologist. New Series, 96*(3), 606-633.

Star Leigh, S. (2010). This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology & Human Values, 35*(5), 601-617. doi:10.1177/0162243910377624

Star Leigh, S., & Griesemer, J. R. (1989). Institutional ecology, "translations" and boundary objects: amateurs and professionals in Berkeley's Museum of Vertebrate Zoology, 1907–39. *Social Studies of Science, 19*(3), 387–420.

Wróblewska, M. N. (2017a). Ewaluacja 'wpływu społecznego'? Nie kopiujcie Brytyjczyków! Wywiad z Davidem Sweeney, dyrektorem HEFCE ds. badań, edukacji i transferu wiedzy. *Nauka i Szkolnicwo Wyższe, 1*(49), 157-166. doi:http://dx.doi.org/10.14746/nisw.2017.1.8

Wróblewska, M. N. (2017b). Ewaluacja „wpływu społecznego" nauki. Przykład REF 2014 a kontekst polski. *Nauka i Szkolnicwo Wyższe, 1*(49), 79-104. doi:http://dx.doi.org/10.14746/nisw.2017.1.5