



Universiteit
Leiden
The Netherlands

Towards evaluating the research impact made by universities of applied sciences

Coombs, S.K.; Meijer, I.

Citation

Coombs, S. K., & Meijer, I. (2021). Towards evaluating the research impact made by universities of applied sciences. *Science And Public Policy*, 48(2), 226-234.
doi:10.1093/scipol/scab009


Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](#)

Downloaded from: <https://hdl.handle.net/1887/3279350>

Note: To cite this publication please use the final published version (if applicable).

Towards Evaluating the Research Impact made by Universities of Applied Sciences

Sarah K. Coombs ^{1,*} and Ingeborg Meijer²

¹Library, Saxion University of Applied Sciences and External PhD Candidate Centre for Science and Technology Studies, University Leiden, P.O. Box 70.000 Enschede, 7500 KB, The Netherlands and ²Centre for Science and Technology Studies, Leiden University, P.O. Box 905 Leiden, 2300 AX, The Netherlands

*Corresponding author. E-mail: s.k.coombs@cwts.leidenuniv.nl

Abstract

Given the mandate of Universities of Applied Sciences (UASs) to create an impact on society, the evaluation of their research impact is of great importance. And yet, the methodology for evaluating this impact appear less explicitly in research literature than other forms of research. The purpose of this article is to present a literature-based analysis to discover from the complex world of existing theories and frameworks what criteria, assumptions and requirements are relevant for evaluating the impact of applied research. This article will also discuss the relevancy of frameworks currently used for research impact evaluation and the potential they have for operationalising, enriching and supporting the current national evaluation framework used by Dutch UASs. Finally, this article will conclude that the recommendations necessitate the creation of a new framework where the context and process of practice-based research and their stakeholders are included.

Key words: research impact evaluation; research impact Frameworks (RIFs); Universities of Applied Sciences; practice-based research; applied research

A binary higher educational system is one in which a distinction is made between academic universities and other higher educational institutions (Kyvik and Lepori 2010). Several European countries including the Netherlands maintain this system known by names such as Technikons, polytechnics, Fachhochschulen and hogeschole. These *Universities of Applied Sciences* (UASs) deliver a highly trained workforce that is innovative and knowledgeable about research that supports or enhances innovation (Jongbloed 2010). They fulfill the triple role of a UAS, which is to: educate; connect to industry and society; and do research that facilitates these endeavours (Kyvik and Lepori 2010).

The nature of research conducted by UASs, applied oriented and practice based, often differs from research done at traditional Universities (Kyvik and Lepori 2010). The requirements for evaluating the societal impact of this type of research may, therefore, also differ. Given the nature of applied sciences, to conduct problem-oriented research that originated in society, and the mandate of UASs to create an impact on society, the evaluation of research impact is perhaps more important for the applied sciences (Kyvik and Lepori 2010). Yet how the evaluation of such research should be accomplished appears less explicitly in literature. As a recent article by Pedersen et al. (2020) illustrates, there is a wealth of frameworks, theoretical assumptions, contexts for research impact evaluation but what is required and applicable to the research done by UASs is less well recognised.

The purpose of this article is to present a literature-based analysis to discover from the complex world of theories and frameworks what criteria, assumptions and requirements are relevant for evaluating the impact of applied research. This article will also discuss the relevancy of currently used frameworks for research impact evaluation and the potential they have for operationalizing, enriching and supporting the current national evaluation framework used by the Netherlands Association of Universities of Applied Sciences (NAUAS) known as the BranchprotocolKwaliteitszorgOnderzoek 2016–22 (BKO). Based on the analysis, this article will include recommendations necessary for creating a framework suitable for evaluating practice-based research at UASs.

While scholarly research in Dutch UASs has been a part of their mandate for <20 years (van Gageldonk 2018) the purpose of its research is clear. Since their inception, the role of a UAS has been to influence the world by training future generations to improve, innovate and enhance the development of professions and society (van Gageldonk 2018). This original goal of training students for real world professions rendered the function of conducting research secondary to the development of training capacities. The emphasis was on teaching students the newest techniques and theories that they could then apply to the professions for which they are trained (van Gageldonk 2018). In the last two decades, however, there has been a transition within Dutch UASs as research has been elevated to an accepted component of its core functions in combination with teaching (de Weert and Beerkens-Soo 2009).

UAS research then is to focus on practical applicability, be demand driven and applied to changes within society, be collaborative and multidisciplinary, and, connect to education by incorporating the results into curricula (UAS4Europa 2017). This is accomplished in two ways: through research that is initiated for the development of regional needs; and, through research that strives to improve education and professional practice. By doing so, UASs return to their initial mandate, that is, to educate students for professional careers (Kyvik and Lepori 2010).

These characteristics of University of Applied Sciences research fit into what Gibbon and colleagues' call Mode 2 research (Gibbon et al. 1994; de Weert and Leijnse 2010) as well as Stokes' Pasteur's Quadrant where applied science is recognised as Edison's Quadrants (Stokes 1997; Kyvik and Lepori 2010).

1. The policy

In its publication '*Onderzoek met Impact*' (2016) ('Research with Impact'), the NAUAS outlined a strategic agenda for its 2016–20 research program. This document describes the ten areas of society in which Dutch UASs aim to collectively have impact. Reflecting both the European Commission's Grand Challenges and the United Nations' Sustainable Development Goals, these areas include: Health and Vitality; Education and Talent Development; Resilient Society in Community, City and Region; Smart Technology and Materials; The Built Environment: Sustainable and Liveable; Sustainable Transport and Intelligent Logistic; Sustainable Agriculture; Water and Food Supply; Energy and Energy Supply; Art and Creative Industries; and Responsible and Innovative Business (NAUAS 2015).

The intent of the 'Research with Impact' document clearly illustrates the NAUASs increased concern with impacting society through UAS research. In a follow up publication, '*Meer Waarde in het hbo*' (2018) ('More Value in Higher Professional Education') the NAUAS states the need for an evaluation and monitoring framework that would recognise the impact of research done by Dutch UASs. Such a framework would enable the NAUAS and the UASs they represent to evaluate the extent to which they are fulfilling their impact responsibility (Franken et al. 2018). It would also help to determine if a differentiation exists between policy and practice. This document does not, however, include a means of evaluating the success of the research in impacting society but rather requests that an appropriate evaluation and monitoring framework with practical applicability be found or created (Franken et al. 2018). This underlines the necessity and immediacy of developing such a framework.

2. The BKO

Dutch UASs are not completely without evaluation. The national evaluation framework currently used, BKO 2016–22, is an ex-post general evaluation approach used to provide the NAUAS with an all-encompassing evaluation of a *lectoraat* (research group) (van Drooge 2016). The NAUAS refers to it as '*Kwaliteitszorgstelsel*' (quality assurance system) for the maintenance and bettering of the quality of practice-based research; how it is organised, and the organisations supporting it (NAUAS 2015). The current version spans from 2016 to 2022 and is the second version (van Gageldonk 2018). It was developed in parallel with the well-known SEP protocol (KNAW, VSNU, NOW 2016). It consists of five criteria: research group vision and indicators to express this; organisation of

the group including people power, finances, internal/external partnerships, networks and relationships; research quality; relevance and impact on: professional practice and society, education and professionalization, knowledge development within the research domain; regular and systematic evaluation of research process and results. This evaluation takes place every six years and includes experts, peers and stakeholders in the evaluation committee. These evaluations are not centrally archived nor are they openly shared.

According to the BKO, the evaluation of relevance and impact on professional practice and society, education and professionalization, and knowledge development within the research domain, UASs are asked to choose indicators that reflect the following three components of practice based research: research contributes practical knowledge for the professional field and society at large and thereby contributes to innovation; research contributes practical knowledge whereby UASs education remains current and the professionalization of teacher; research contributes to knowledge development. While examples of indicators are given, UASs are responsible for selecting their own. At this time, a critical reflection (narrative) including strong and weak points/characteristics, measures taken for improvement in accordance with the previous evaluation, introduction and accountability for the self-reflection with respect to approach, method, stakeholders and a conclusion on strengths weakness, improvement measures, priorities for the future is requested of the evaluated research group. This critical reflection can also be used to support qualitative indicators where use and impact are included. (For monitoring purposes, UASs are required to annually report on research budget and personnel to the NAUAS.) While adjusted for UAS implementation, the SEP protocol served as the starting point for the BKO and mirrors its format (KNAW, VSNU, NOW 2016).

The current agenda of the NAUAS states that creating impact in society in ten research domains is its priority but is not explicit on how to accomplish this. The BKO contains components often used in a research impact evaluation framework (indicators, narrative) but is far broader than an impact evaluation framework. While stating that context matters, it is not explicit, and there is little guidance in the operationalisation. UASs of the Netherlands require an impact evaluation framework that provides a solid evaluation of the impact their research projects are generating that supports the ex-post BKO evaluations of research group occurring every 6 years. The question then arises as to how can societal impacts be evaluated in the context of the goals of UASs? What is required to accomplish this? And what has already been done that can be applied to evaluating the societal impact of research done by UASs? The following section sets out the theoretical requirements for reaching these goals.

3. Theoretical requirements for evaluating the research impact of UAS research

3.1 Recommended philosophical assumptions

The need to accurately and comprehensively evaluate the societal impact of research is not strictly a UAS problem—it is a very relevant problem for all institutions participating in research similar to the UASs (Bölling and Eriksson 2016). This research is referred to, among other things as, Applied, Triple Helix, Third Mission, Entrepreneurial, Mode 2 or Edison's Quadrant research (Bornmann 2013). In addition, it can overlap with research conducted by traditional universities (de Weert 2011). Nevertheless, pinpointing the specific requirements for evaluating the societal impact of research

done by UASs has proven difficult. Raftery and colleagues address this issue directly in their systematic review where they state that evaluating the research impact of Mode 2 research is best suited to a methodology created from a *realist or performative philosophical assumption* (Raftery et al. 2016).

Often an evaluation approach is based on ‘philosophical assumptions’ made regarding the links between research and societal impact. They include assumptions about ‘the nature of research knowledge, the purpose of research, the definition of research quality, the role of values in research and its implementation, the mechanisms by which impact is achieved, and the implications for how impact is measured’ (Greenhalgh et al. 2016: 2). These assumptions relate to the area of research and help to form and enhance the methods and tools used. These philosophical assumptions include positivist, constructivist, critical, performative, and realist assumptions (Raftery et al. 2016).

3.2 Recommendation one: Realist evaluation

According to Raftery and colleagues, an impact evaluation done from a realist philosophical assumption must consider the different means through which knowledge is taken up and research is used, based on a Context-Mechanism-Output-Impact configuration. Within this realist evaluation, frameworks with a realist philosophical assumption consider the mechanism through which the impact is made and make common assumptions about what works for whom under what conditions (Raftery et al. 2016). Initially introduced by Pawson and Tilley, realist evaluation suggests that research creates output only in so far as they introduce appropriate ideas and opportunities (mechanisms) in the appropriate settings (context) (Pawson and Tilley 1997). Realist evaluation ‘elaborates how mechanisms could work in a given context and asks the people who could know about it to provide evidence’ (Stame 2004: 62). The presupposed mechanism for impact with a realist philosophical assumption is the interaction between the people involved and the resources available for the implementation of findings (Greenhalgh et al. 2016).

Raftery and colleagues’ recommendation of a context driven methodology is of particular importance for research done by UASs. Context determines the operationalization of the concept of societal impact, and, thus, context is essential for creating an applicable evaluation approach. In order to understand the context-mechanism-output, realist evaluation requires the contribution of the ‘people who know’ (Stame 2004: 62). It is assumed in a realist evaluation that the mechanism through which impact is achieved is the interaction between the reasoning of policy makers and practitioners, and the resources available for implementing the findings (Raftery et al. 2016). The stakeholders, in their various forms, who contribute to UAS research, must, therefore, be a part of the evaluation process.

3.3 Recommendation two: Performative assumption

Raftery et al. (2016) also suggest that *performative assumption* is possible. According to Greenhalgh et al. (2016), a performative assumption relies on Actor-Network Theory to focus on the connection established between people and technology that lead to the creation of new entities. In order for research to have an impact, a realignment of actors, human or technological, must occur. Thus, a societal impact evaluation with a performative assumption must ‘focus on the changing actor scenario and how this gets stabilised in the network’ (Greenhalgh et al. 2016: 3). Frameworks from this

philosophical assumption assume impact mechanisms are changes in the actor-network that occur through the creation of new configurations between actors. These changes come about as a result of both formal and informal interactions. Societal impact evaluations based on a performative assumption thus take the process of impact creation into account and attempt to map these interactions and changes (Raftery et al. 2016).

3.4 Recommendation three: Co-production model

Raftery and colleagues suggest that an impact evaluation from a performative assumption should be accompanied by a *co-production model* (2016). They go so far as to say that it can in fact be referred to as a co-production model (Raftery et al. 2016). Initiated in the 1970s by Elinor Ostrom, co-production models stress the need for contribution from stakeholders throughout the creation process including planning, designing, delivery, and auditing of the service (Boyle et al. 2006). Further, there is an expectation that through their contribution to the creation of the service, in this case the evaluation, stakeholder contribution will create synergy between the various people and groups involved (Branden and Pestoff 2006). The use of a co-production model also assumes a long-term perspective for the results. Creation of a co-production model often results in stakeholders experiencing a shared responsibility for the outcomes. A true co-production model results in a shift in power whereby the stakeholders take the lead from the evaluator and take responsibility for the outcome (Bovaird 2007). Ramaswamy and Ozcan (2014) have suggested, in order for this to occur, stakeholders must see the value of the process and outcome. This is best created by focusing on the stakeholder experiences and giving stakeholders the opportunity to interact with each other face to face.

However, recent work by Oliver et al. (2019) suggests that although this type of research practice is often recommended, it is not without its challenges. Co-production requires personal interaction and all the inherent challenges that human nature brings. These challenges include disagreements within the stakeholder groups, pressure to produce certain outcomes or omit certain results, being ‘too helpful’ with analysis and resources, thus creating the potential for bias and other scientifically questionable results. Each of these challenges results in costs, be it financial, temporal, relational, reputational, or ethical. Therefore, the advantages and disadvantages should be weighed before embarking on this type of process (Oliver et al. 2019). Nevertheless, because of stakeholder inclusion, the results of co-production research are often ready for implementation earlier than other models because needs, capacities and priorities have already been taken into account (Oliver et al. 2019). The effect created by including the stakeholders in the process suggests that the very nature of the recommended methods for creating a usable evaluation approach for Mode 2 research initiates the adoption process (Adam et al. 2018).

3.5 Recommendations four and five: Formative and ‘real time’ evaluation

The recommendation of a co-production model is further supported by recent work done by van Drooge and Spaapen (2017). They suggest, like Raftery and colleagues, that Mode 2 research should involve *formative evaluation*. They also suggest that trans-disciplinary research requires formative evaluation ‘where learning is the prime motive for evaluation, the focus is on the variegated context in which research and innovation takes place’ (van Drooge and Spaapen 2017: 2). They, too, stress the need for context of

application to be considered when evaluating Mode 2 research, as well as stakeholder inclusion to create a joint responsibility between participants where ‘mutual learning and improving the research effort’ is central for improving the research impact of Mode 2 research (van Drooge and Spaapen 2017: 6). By using a bottom-up approach, accountability for impacting society is no longer something to be assessed through ex post means—it is assumed. Because society has been included in the research, the question becomes not *if* society has been impacted but *how* society has been impacted and *how* it can be further impacted in the future (van Drooge and Spaapen 2017).

Raftery and colleagues state that an impact evaluation of Mode 2 research should be *formative* and in *real-time*, and, take the ‘messy, unpredictable and evolving interaction’ into account (Raftery et al. 2016). In the RAND publication ‘Measuring research: A guide to research evaluation frameworks and tools’, Guthrie et al. (2013) agree with this, suggesting that formative evaluation compliments the characteristics of Mode 2 research.

More specifically, Guthrie et al. (2013) state that a formative societal impact evaluation of cross or multidisciplinary research should utilize case studies, document review and peer-review as tools for accomplishing this. Raftery et al. (2016) also state that in-depth case studies are required for understanding the shifting nature of applied sciences. According to Greenhalgh et al. (2016), current evaluation frameworks for evaluating societal impact frequently consist of three parts: case studies for explaining the process and interactions that come as a result of knowledge production impacting society; a narrative required for explaining the feedback loops and non-linear nature of impact, as well as why certain outcomes expected to make impact fail; and a logic model which is a visualisation of the input activities and output and outcomes of impact (Greenhalgh et al. 2016).

The authors of these publications appear to agree that the requirements for evaluating the societal impact of research done by UASs are formative, real-time evaluation, where stakeholders are included to create a bottom-up approach for research. They also agree on the use of the case study as a tool for formative research evaluation. However, these experts do not necessarily agree on the use of a logic model.

3.6 Recommendation six: Logic models

Guthrie et al. (2013) present a neutral stance on the subject of logic models. They suggest that the logic model, like data visualisation, is a tool that can be used for any type of societal impact evaluation. Raftery et al. (2016), however, are quite passionate about the use of a logic model. They suggest that many evaluation frameworks utilise a positivist logic model as one of their tools to illustrate how:

‘causal connections in the temporal sequence of inputs (research funding), process (execution of discrete projects or programmes of research, usually following a predefined protocol), outputs (e.g. publications and presentations) and outcomes (impacts on end-users of research), the study of knowledge production has emphasised the non-linearity, messiness and unpredictability of the collaborative knowledge production process’ (Raftery 2016: 59).

However, the ‘collaborative knowledge production process’ in Mode 2 knowledge production is created through application (Raftery et al. 2016: 59). Raftery et al. (2016) suggest that an approach including a logic model is inadequate for Mode 2 research because of the complex levels of interactions that occur in Mode 2

research. This study goes on to say that most Mode 1 research can be effectively evaluated with a logic model but that attempting to squeeze Mode 2 research into these types of frameworks does not do it justice. They further suggest that a logic model is in fact a tool primarily utilised by evaluations with a positivist philosophical assumption where knowledge is seen as fixed and stable (Greenhalgh et al. 2016). Thus, the presence of a logic model in a framework implies it is not suitable for evaluating the research impact of Mode 2 research.

Based on these considerations, it can then be suggested that the recommendations for requirements when evaluating research done by UASs include:

- a realist philosophical assumption where evaluation is based on context-mechanism-output, or;
- a performative philosophical assumption in which knowledge is a process; and
- a co-production model; and
- a focus on formative, ‘real-time’ evaluation; and,
- no reliance on an existing logic model.

3.7 What existing methods can be applied to evaluating the impact of applied research?

It is against this backdrop of requirements that current models can be reviewed for applicability. Recent work by Adam et al. (2018) suggests that the use of a conceptual framework is important for the simplification of research impact evaluation. Frameworks also increase comparability and communication over the results. The use of a framework also assists in addressing hurdles frequently encountered when striving to evaluate impact. These methodological issues include ‘attribution (assigning the right impact to a specific piece of research or vice versa), time-lag (determining the time for impact and the right timing to engage in research impact assessment), and the counterfactual (examining what would have happened if the given piece of research did not occur)’ (Adam et al. 2018: 9). Table 1 provides an overview of established frameworks and assesses how suitable they are for evaluating the research impact of UASs according to the requirements stated earlier.

3.8 Comparison of frameworks for UAS research

In a study comparing RIFs conducted by Greenhalgh et al. (2016), more than 20 existing models and frameworks for research impact evaluation were referenced. Of those original 20, six approaches were repeatedly referenced. These include the Payback Framework and two of its derivatives, RIF and Canadian Academy of Health Science Framework (CAHS). The Payback Framework has been used as a starting point for more than 40 other approaches for evaluation but in addition to Payback itself, RIF and CAHS are the most frequently cited (Raftery et al. 2016). Also included are Monetisation, the UK Research Excellence Framework (REF), and Societal Impact Assessment (SIA). Two well-known frameworks ERiC and SIAMPI fall under the heading of SIA. Several of the same authors were involved in a little-known Dutch research evaluation guideline known as Waardevol (Valuable) (van Drooge et al. 2011). This too falls under the heading of SIA.

Greenhalgh et al. (2016) suggest that as a consequence of their consistent reference, international influence, and impact on policy, the above mentioned six approaches, Payback, RIF, CAHS, Monetisation frameworks, REF, and SIA can be considered established approaches for measuring research impact. Because of its

Table 1. Overview of the suitability of established frameworks.

Main framework	Derivative framework	Co-production model	Philosophical assumption	Formative/summative	Ex ante/Ex post/real-time	Logic model
Payback	Payback (Buxton and Hanney 1996, Donovan and Hanney 2011)		Positivist	Summative	Ex post	✓ (Raftery et al. 2016; Greenhalgh et al. 2016)
	CAHS (Canadian Academy of Health Sciences 2009)		Positivist	Summative	Ex post	✓ (Raftery et al. 2016; Greenhalgh et al. 2016)
	RIF (Kuruville et al. 2006; Kuruville et al. 2007)		Positivist	Summative	Ex post	✓ (Raftery et al. 2016; Greenhalgh et al. 2016)
SIA	ERiC (van Drooge 2007; ERiC 2010)		Performative and constructivist		Ex ante or expost	
	SIAMPI (van Droog and Spaapen et al. 2017, Propp 2010)		Performative and constructivist	Formative or summative (Penfield et al., 2013)	Ex Ante or Expost	
	Waardevol (van Drooge et al. 2011)		Performative and constructivist			
Monetisation (Raftery et al. 2016; Greenhalgh et al. 2016)						✓ (Raftery et al. 2016; Greenhalgh et al. 2016)
REF (HEFCE 2012)						
Contribution Mapping (Kok and Schuit 2012)			Performative	Summative of formative	Real-time	
ASIRPA (Joly et al. 2015, Matt et al. 2017)			Realist	Formative	Ex Post	
PIPA (van Droog and Spaapen 2017)		Co-production	Realist and performative	Formative	Real-time	✓ (van Drooge and Spaapen 2017)
PRO (van Beest, Baljé and Andriessen 2017)						✓ (van Beest, Baljé and Andriessen 2017)

innovation, Contribution Mapping introduced by Kok and Schuit (2012) was also included in their study. Not viewed as an established framework, this approach can be seen as a variation on SIA with different authors and a noticeable shift of philosophical assumption (Greenhalgh et al. 2016).

Also included above is the ASIRPA framework. This framework was developed in the context of an agricultural impact project to develop an international methodological standard for assessing societal impact (Joly et al. 2015). There is currently no evaluation model available for or from UASs themselves, other than the general BKO. However, based on the Technology Readiness Levels model, the Praktijkgeredheid van Onderzoek (Practical Readiness of Research) (PRO) model by van Beest et al. (2017) strives to provide researchers with a tool that can be used regardless of the research domain. While it appears as a logic model, the PRO-Model strives to aid in: identifying research goals and connected activities to be pursued in this project; assessing which research activities are to be left for others; and, identifying in which order previously selected goals are to be pursued for the creation of change. This approach encourages discussion over the practical relevancy and methodological grounding of UAS research (van Beest et al. 2017: 53). For the sake of completeness, the PRO-Model has been included in the comparison presented in Table 1. Finally, the evaluation and monitoring system Participatory Impact Pathways Analysis (PIPA) as executed by van

Drooge and Spaapen (2017) has also been included. This process driven evaluation and monitoring system strives to evaluate the societal impact of transdisciplinary research.

3.9 What fits?

It can be concluded from the above table that there is no perfect fit between the established frameworks and the proposed requirements for evaluating the societal impact of research done by UASs. The majority of the described approaches are created from positivist and constructivist assumptions (Raftery et al. 2016). None of the established examples mentioned are co-production and many of them are summative instead of formative. Also, many of these established frameworks utilise a preconceived positivist logic model as one of their tools (Greenhalgh et al. 2016) that does not take the nature of Mode 2 research into account (Raftery et al. 2016). However, as Table 1 also indicates, there are three frameworks that fulfill parts of the recommended requirements that can act as possible starting places. These include ASIRPA, Contribution Mapping, and the PIPA evaluation and monitoring system van Drooge and Spaapen.

Although increasing in number, examples of realist evaluations, and co-production in impact evaluation are few (Raftery et al. 2016). ASIPRA is a theory-based realist evaluation that makes use of contribution and productive interaction to help assess long-term

impact (Joly et al. 2015). While creating it, the authors also took Payback, the most cited framework to date (Greenhalgh et al. 2016), into account (Joly et al. 2015). What makes ASIRPA stand out is its attempt to create a useable framework in practice through the use of standardised case studies that combine quantitative and qualitative methodologies that can be used over a range of disciplines, are comparable, and can be aggregated (Joly et al. 2015). Stressing the need for context-mechanism-impact, this framework utilises a set of tools including chronology and vector of impacts. It uses PIPA first introduced by the Consultative Group on International Agricultural Research (CGIAR) that stress the non-linearity of impact and the need for stakeholder contribution to the generation of impact. However, ASIRPA is currently an ex-post framework and falls short in the real-time and co-production areas. Although stakeholders are interviewed, and networks and stakeholders are taken into account, there is no concrete co-production component to this framework. The inclusion of stakeholders from the onset in the creation of the evaluation process is essential for creating an approach that can be used for the applied sciences (Greenhalgh et al. 2017). ASIRPA will need to be modified to real-time and be more of a co-production model in order to be fully useable for UAS research use.

Kok and Schuit's (2012) Contribution Mapping also fulfills several of the requirements previously identified. This is clearly a performative, real-time, formative evaluation based on actor-network theory. It focuses on contribution to impact rather than the attribution of the ultimate impact of the research. It uses structured interviews with stakeholders in in-depth case studies to 'map research-related contributions and relate these contributions to alignment efforts' (Kok and Schuit 2012: 2). This three-phase mapping framework focuses on activities and what they refer to as 'alignment efforts' of 'linked actors' and 'key users' that ultimately contribute to the impact of research (Kok and Schuit 2012). By doing so it focuses on process and strives to create 'an account of how the network of actors and artefacts shifts and stabilises (or not)' (Greenhalgh et al. 2016: 11). Although it identifies linked actors and key users, their contribution to the evaluation is limited. The inclusion of stakeholder interviews introduces a co-production component, but like ASIRPA, there is a very limited use of stakeholder contribution and thus a limited concrete co-production component.

van Drooge and Spaapen's (2017) approach, however, has a very intense co-production component. This approach fulfills the real-time, formative, co-production model requirements from a clear realist perspective. Taking the co-production model, a step further than Kok and Schuit, van Drooge and Spaapen (2017) state that stakeholders and evaluators should, in fact, work together to create what they refer to as a logic framework. Using the same impact pathways (PIPA) initiated by the CGIAR mentioned earlier by ASIRPA as a starting point, van Drooge and Spaapen (2017) state that when evaluating transdisciplinary research, a realist 'theory of change' is required. Written as a narrative and taking stakeholders expectations, assumptions, needs and requirements into account this 'theory of change' aims to explain the logical steps, or 'pathways' towards a desired ultimate impact. These are set into a logical framework based on 'inputs, outputs, outcomes and impacts' (van Drooge and Spaapen 2017). From there, the theory of change is strengthened through discussion of possible relationships between the components of the logic framework as well as the 'causal assumption' required to reach the end impacts. By doing this, van Drooge and Spaapen (2017) believe that a 'theory of change opens up this linear narrative and it allows for different contributions coming from

different angles in society to participate in the debate about how to achieve a particular desired change' (van Drooge and Spaapen 2017: 50). This appears then to take the 'collaborative knowledge production process' into account as well as the non-linearity stressed by Raftery et al. (2016). However, this proposed work process is extremely time consuming and consequently not necessarily feasible for regular use (van Drooge and Spaapen 2017).

3.10 A critical reflection on the proposed requirements

Is it then the use of an existing logic model that is the issue rather than one created with stakeholders? Is Raftery and colleagues objection to a logic model in fact an objection to an existing logic model? It appears that a logic model created through co-production may be able to bring the various layers and messiness of Mode 2 research into view. However, it may also bring with it the same preconceptions that occur with the use of an existing logic model. Raftery et al. (2016) also state that the presence of a logic model correlates to methodologies with a positivist philosophical assumption which is not appropriate for Mode 2 research (Raftery et al. 2016). Given this discrepancy, it is preferable to focus on co-production as a paradigm, rather than explaining that logic model use is permitted if it is not preconceived.

One could also argue, however, that the entire bases of a realist evaluation is in and of itself a logic model. The formula of context-mechanism-output-impact could be interpreted as a linear expression of impact creation. This would then lead to the same argument that Mode 2 research cannot be squeezed into the linear confines of a logic model. This leads to the question of whether a realist philosophical assumption that is based on context-mechanism-output-impact is useable for evaluating the research done by UASs.

As the previous analysis shows, the concept of working with a philosophical assumption is confusing. Whereas a realist philosophical assumption is clearly based on a tradition with history, a performative assumption is based on Actor-Network theory and is easily confused with a performance-based evaluation. It is difficult to find corroborating information on performative assumptions.

What each of these requirements share, however, is a focus on the process of impact creation. Be it through context-mechanism-output-impact, actor-network theory, learning through evaluation in real-time, it is the process that stands centrally. It is the research process and thus the process of impact creation that needs to be monitored in order for evaluation to be possible. While the BKO is not currently designed to do this, a theoretically grounded impact evaluation would act to enhance it by describing not only the outcomes but also the process through which research impact is created.

4. Discussion

4.1 From theory and frameworks into operationalization—The inclusion of stakeholders

The stakeholder is central for the operationalization of the requirements for evaluating the impact of research done by UASs. The nature of this research means that a broad range of stakeholders exist in this type of research. In this case, while the direct researcher is the primary stakeholder, the partners they work with must also be included. It is in fact the engagement of non-academic stakeholders that can make this process successful (Adam et al. 2018). These partners come from relationships with industry, government and society as well as the funders that support them (Greenhalgh et al. 2017).

For Dutch UASs, this includes a wide range of groups and organisations; health centres like hospitals and retirement homes, museums, sports clubs, educational institutions, large and small businesses, and industrial partners, to name only a few. All of these stakeholders are potential end users of this evaluation approach at different levels.

A component of the BKO Standard Two requires that the relevance, intensity and sustainability of internal and external partnerships, networks and relationships in people and resources be evaluated with respect to the realisation of the research profile. It also asks for self-reflection on stakeholders in the narrative. While this includes information about the stakeholders, it does not include the participation of stakeholders themselves. The BKO asks for stakeholder participation in the evaluation committee but this is limited to one or two participants. The inclusion of stakeholders in a co-produced impact evaluation is necessary for the insight into the diverse and variegated societal impact of UAS research. It is a necessary theoretically grounded step towards augmenting the function of the general BKO.

The recommendation for a co-production model and a performative or realist evaluation requires that an evaluation of the impact of research done by UASs is based on a bottom-up approach that includes the various stakeholders involved while taking the process of impact creation into consideration. Raftery et al. (2016) suggestion of a context driven methodology is of particular importance for research done by UASs. Context determines the operationalization of the concept of societal impact, and, thus, context is essential for creating an applicable approach. The real-time component of these requirements means that these stakeholders need to be included from the beginning of the process. This too is part of Mode 2 research as the inclusion of the stakeholder from the beginning also helps create more socially robust knowledge that can be effectively translated into practice (Adam et al. 2018).

The contribution of stakeholders is also required at the end of the process when output comes to fruition. Unlike INRA, where a database of information is available for the in-depth, standardised case studies set out by ASIRPA, Dutch UASs lack such a resource. Currently 27 Dutch UASs make use of a shared repository for print output. To date, five Dutch UASs also make use of a CRIS. Three subscribe to a well-known commercially obtained CRIS. Another is preparing for the implementation of a recently revamped system designed by a Dutch University. The last example has designed their own system based on the needs of their researchers, quality control office and other support staff. This institution also has its own repository where research output can be stored regardless of form. However, although they are working towards achieving it, even this system lacks the relevant societal information required for evaluating research impact (van der Graaf 2018; Woertman and Doove 2019).

5. Conclusion

Finding an appropriate means of evaluating the research impact of research done by UASs has proven complicated. Raftery et al. (2016) backed up by van Drooge and Spaapen (2017) and Gutherie et al. (2013) suggest that in order to best evaluate Mode 2 research similar to that achieved in Dutch UASs, a formative real-time evaluation should be used from a realist perspective that includes context-mechanism-output. Or a performative philosophical assumption with a co-production model without making use of a preformulated

logic model. If these recommendations are to be put into place, there is, to date no 'established' framework or approach that is 'cut and paste' ready for use by UASs.

Three frameworks present possible starting points for creating a suitable approach:

- A. ASIRPA provides a realist evaluation that incorporates Participatory Impact Pathways suggested by van Drooge and Spaapen (2017) as well as in-depth case studies, made easier through standardisation for realistic utilisation. It does, however, neglect the co-production and real-time evaluation;
 - B. Contribution Mapping is real-time and formative. This framework provides a performative assumption where the process of impact creation is key. However, stakeholder inclusion is limited to structured interviews; and
- A. vanDrooge and Spaapen's evaluation and monitoring use of PIPA is a formative, real-time, realist evaluation focused on the 'theory of change'. Unfortunately, the use of stakeholders for creating a logic model which is central to the evaluation and monitoring use can become impractical (van Drooge and Spaapen 2017).
 - B. What each of these three frameworks have in common is that the process of impact creation is what is important. And what the recommended components of impact evaluation for research done by UASs suggests is that relevant stakeholders are essential from the beginning of this process. This need for stakeholder inclusion in the process means that regardless of which framework is chosen, a new evaluation approach for UAS research is required. By including the relevant stakeholders, the missing link between the general BKO and the theoretical foundations is bridged.

Given the short history of research conducted by UASs, it is not surprising that there is no recognised approach for evaluating the impact created by the research of these institutions. The nature of Mode 2 or Edison's quadrant research puts research impact at the heart of its mission and the process whereby the research impact is produced. In order to assess if the Netherlands Association of UASs has accomplished their goal of creating impact in society in their research domains, an evaluation approach for research impact is required. This approach will assist in providing insight into the impact of research carried out by researchers of Dutch UASs. The motivation for this is found in the research task of UASs, to conduct research that stems from a challenge in society. As challenges change and research done by UASs continues to mature it is increasingly important that it appropriately conveys the impact it is creating.

Conflict of interest statement. None declared.

REFERENCES

- Adam, P., Ovseiko, P., Grant, J., et al. (2018) 'ISRIASatement: Ten-Point Guidelines for an Effective Process of Research Impact Assessment', *Health Research Policy and Systems*, 16/1: 8.
- Association of Universities in the Netherlands (VSNU), the Netherlands Organisation for Scientific Research (NWO), and the Royal Netherlands Academy of Arts and Sciences (KNAW). (2016) *Standard Evaluation Protocol 2015-2021*.
- Bölling, M. and Eriksson, Y. (2016) 'Collaboration with Society: The Future Role of Universities? Identifying Challenges for Evaluation', *Research Evaluation*, 25/2: 209–18.

- Bornmann, L. (2013) 'What is Societal Impact of Research and How Can It be Assessed? A Literature Survey', *Journal of the American Society for Information Science and Technology*, 64/2: 217–33.
- Bovaird, T. (2007) 'Beyond Engagement and Participation: User and Community Co-production of Public Services', *Public Administration Review*, 6/7: 846–60.
- Boyle, D., Clarke, S. and Burns, S. (2006) *Aspects of co-Production: The Implications for Working, Health and Volunteering*. New Economics Foundation: London.
- Branden, T. and Pestoff, V. (2006) 'Co-production, the Third Sector and Delivery of Public Services', *Public Management Review*, 8/4: 493–501.
- Budtz Pedersen, D., Følsgaard Grønvald, J. and Hvidtfeldt, R. (2020) 'Methods for Mapping the Impact of Social Sciences and Humanities—A Literature Review', *Research Evaluation*, 29/1: 4–21.
- Buxton, M. and Hanney, S. (1996) 'How Can Payback from Health Services Research Be Assessed?', *Journal of Health Services Research & Policy*, 1/1: 35–43.
- Canadian Academy of Health Services. (2009) *Making an Impact*, pp.1–143. The Canadian Academy of Health Services: Ottawa.
- de Weert, E. (2011) 'Transformation of System Convergence?'. In: J., Enders, H., de Boer and D., Westerheijden (eds.) *Reform of Higher Education in Europe*, pp.103–22. Sense Publishers: Rotterdam.
- and Beerkens-Soo, M. (2009) *Research at Universities of Applied Sciences in Europe, Conditions, Achievements and Perspectives, on the Initiative of the European Network for Universities of Applied Sciences. (European Project: Educating the New European Professional in the Knowledge Society (EDUPROF))*. European Network for Universities of Applied Sciences (UAS): The Hague.
- and Leijnse, F. (2010) 'Practice-Oriented Research: The Extended Function of Dutch Universities of Applied Sciences'. In: S., Kyvik and B., Lepori (eds.) *The Research Mission of Higher Educational Institutions outside the University Sector: Striving for Differentiation*, pp.199–217. Springer Science+Business Media: London.
- Donovan, C. and Hanney, S. (2011) 'The 'Payback Framework' Explained', *Research Evaluation*, 20/3: 181–3.
- Franken, A., Andriessen, D., van der Zwan, F., et al. (2018) *Meer Waarde Met Hbo: Doorwerking Praktijkgericht Onderzoek Van Het Hoger Beroepsopleidings. Vereniging Hogescholen: The Hague.* <https://www.vereniginghogescholen.nl/system/knowledge_base/attachments/files/000/000/961/original/RAPPORT_MEER_WAARDE_MET_HBO.pdf?1537795313> accessed 27 November 2019.
- Gibbon, M., Limoges, C., Nowotny, H., et al. (1994) *The New Production of Knowledge: Dynamics of Science and Research in Contemporary Societies*. Sage Publications: London.
- , Nowotny, H. and Scott, P. (2001) 'Mode 2' Revisited: The New Production of Knowledge', *Minerva*, 41: 179–94.
- Greenhalgh, T., Ovsieiko, P., Fahy, N., et al. (2017) 'Maximising Value from a United Kingdom Biomedical Research Centre: Study Protocol', *Health Research Policy and Systems*, 15/1: 70.
- , Raftery, J., Hanney, S., et al. (2016) 'Research Impact: A Narrative Review', *BMC Medicine*, 14/1: 78.
- Gutherie, S., Wamae, W., Diepeveen, S., et al. (2013) *Measuring Research: A Guide to Research Evaluation Frameworks and Tools*, pp.1–171. RAND Corporation: Arlington, VA.
- HEFCE. (2012) Higher Education Funding Council of England: 2014 REF: Assessment Framework and Guidance on Submissions. *Panel A criteria*. HEFCE: London.
- Joly, P., Gaunand, A., Colinet, L., et al. (2015) *ASIRPA: A Comprehensive Theory-Based Approach to Assessing the Societal Impact of a Research Organizations*. Institut National de la Recherche Agronomique: Grenoble.
- Jongbloed, B. (2010) 'The Regional Relevance of Research in Universities of Applied Sciences'. In: S., Kyvik and B., Lepori (eds.) *The Research Mission of Higher Educational Institutions outside the University Sector*, pp. 3–21. Springer Science+Business Media: Berlin, Germany.
- Kok, M. and Schuit, A. (2012) 'Contribution Mapping: A Method for Mapping the Contribution of Research to Enhance its Impact', *Health Research Policy and Systems/Systems*, 10: 21.
- Kuruville, S., Mays, N., Pleasant, A., et al. (2006) 'Describing the Impact of Health Research: A Research Impact Framework', *BMC Health Services Research*, 6:134.
- , — and Walt, G. (2007) 'Describing the Impact of Health Services and Policy Research', *Journal of Health Services Research & Policy*, 12/Suppl 1: 23–31.
- Kyvik, S. and Lepori, B., eds (2010) *The Research Mission of Higher Educational Institutions outside the University Sector*, pp.3–21. Springer Science+Business Media: Berlin, Germany.
- Lepori, B. and Kyvik, S. (2010) 'Sitting in the Middle: Tensions and Dynamics of Research in UASs'. In: S. Kyvik and B., Lepori (eds.) *The Research Mission of Higher Educational Institutions outside the University Sector*, pp. 259–71. Springer Science+Business Media: Berlin, Germany.
- Matt, M., Gaunand, A., Joly, P., et al. (2017) 'Opening the black box of impact – Ideal type impact pathways in a public agricultural research organization', *Research Policy*, 46/1: 207–18.
- Netherlands Association of Universities of Applied Sciences. (2015) *Brancheprotocol Kwaliteitszorg Onderzoek*, pp. 1–26. Vereniging Hogescholen: The Hague.
- (2016) *Onderzoek Met Impact: Strategische Onderzoeksagenda 2016–2020*, pp. 1–40. Vereniging Hogescholen: The Hague.
- Netherlands Code of Conduct for Research Integrity. (2018) 10.17026/dans-2cj-nvwu.
- Oliver, K., Kothari, A. and Mays, N. (2019) 'The Dark Side of Coproduction: Do the Costs Outweigh the Benefits for Health Research?', *Health Research Policy and Systems*, 17/33: 1–10.
- Pawson, R. and Tilley, N. (1997) *Realist Evaluation*. Sage Publishing: London.
- Penfield, T., Baker, M., Scoble, R., et al. (2013) 'Assessment, Evaluations, and Definitions of Research Impact: A Review', *Research Evaluation*, 23/1: 21–32.
- Propp, T. (2010) *Social Impact Assessment Methods for Research and Funding Instruments through the Study of Productive Interactions between Science and Society (SIAMPI). Work Package 3: Nanosciences and Nanotechnology (NST). Case Study: Netherlands. Final Report: 'Conditions, Forms and Evaluation Tools of Knowledge Uptake from Potentially Enabling Nanoscience and Technology (Nanolithography, Nanoelectronics and Nanomaterials).'* pp. 1–115. Rathenau Institute: The Hague.
- Raftery, J., Hanney, S., Greenhalgh, T., et al. (2016) 'Models and Applications for Measuring the Impact of Health Research: Update of a Systematic Review for the Health Technology Assessment programme', *Health Technology Assessment*, 20/76: 1–254.
- Ramaswamy, V. and Ozcan, K. (2014) *The Co-Creation Paradigm*. Stanford University Press: Stanford.
- Spaapen, J., van Drooge, L., Propp, T., et al. (2013) *SIAMPI Final Report: Societal Impact Assessment Methods for Research and Funding Instruments through the Study of Productive Interactions between Science and Society*, pp. 1–36. SIAMPI: Brussels.
- Stame, N. (2004) 'Theory-Based Evaluation and Types of Complexity', *Evaluation*, 10/1: 58–76.
- Stokes, D. (1997) *Pasteur's Quadrant: Basic Science and Technological Innovation*, 1st ed., pp.vii–180. The Brookings Institution: Washington, D.C.
- vanBeest, W., Baljé, J. and Andriessen, D. (2017) 'Hoe Meet je de Praktische Relevantie: Nieuw Model voor Praktijkgerichte Onderzoekers', *Thema*, 17/4: 52–7.
- van Droog, L. (2016) *Valorisation [Ebook]*. Rathenau Instituut: The Hague
- van Drooge, L. (2007) *Measure! The Societal Quality of Research: The Sci_Quest Method Developed and Put into Practice*. ERIC, Evaluation in Context: The Hague.
- and Spaapen, J. (2017) 'Evaluation and Monitoring of Transdisciplinary Collaborations', *The Journal of Technology Transfer*, <http://doi.org/10.1007/s10961-017-9607-7>.
- , Vandenberg, R., Zuidam, F., et al. (2011) *Waardevol: Indicatoren Voor Valorisatie*, pp. 1–35. Rathenau Institute : The Hague.

- van Gageldonk, K. (2018) *Op Zoek Naar Onderzoek: De Totstandkoming Van de Onderzoekfunctie in Het Hoger Beroepsonderwijs*. Universiteit Leiden:Leiden.
- van der Graaf, M. (2018) *Naar Een Betere Zichtbaarheid Van Praktijkgericht Onderzoek Aan de Hand Van 4 User Stories*. Pleiade Management and Consultancy: Amsterdam.
- Drooge, L., van den Besselaar, P., Elsen, G. M. F., et al. (2010). *Evaluating the societal relevance of academic research: A guide*. (ERIC). ERiC-Evaluating Research in Context: The Hague. <http://www.rathenau.nl/en/themes/project/eric-evaluating-research-in-context.html>, Accessed 23 Oct 2018.
- Woertman, E. and Doove, J. (2019) *Nationaal Platform Praktijkgerichtonderzoek*. SURF: Utrecht.