Research Evaluation, 30(4), 2021, 500–513
doi: 10.1093/reseval/rvab026
Advance Access Publication Date: 1 October 2021
Article

OXFORD

Self-evaluating participatory research projects: A content validation of the InSPIRES online impact evaluation tool

Florence Gignac[1,2], Anne-Sophie Gresle[1], Valeria Santoro Lamelas[1,3,*], Montserrat Yepes-Baldó[4], Leonardo de la Torre[1] and Maria-Jesus Pinazo[1], and on behalf of the InSPIRES Consortium

[1]Barcelona Institute for Global Health (ISGlobal), Hospital Clínic-Universitat de Barcelona, Rosselló 132, Barcelona 08036, Spain, [2]Department of Experimental and Health Sciences, Universitat Pompeu Fabra (UPF), Doctor Aiguader, 88, Barcelona 08003, [3]Department of Social Psychology and Quantitative Psychology, Interaction and Social Change Research Group (GRICS), Universitat de Barcelona, Campus Mundet, Passeig de la Vall d'Hebron, 171, Barcelona 08035, Spain and [4]Department of Social Psychology and Quantitative Psychology, Research Group in Social, Environmental and Organizational Psychology (PsicoSAO), Universitat de Barcelona, Campus Mundet. Passeig de la Vall d'Hebron, 171, Barcelona 08035, Spain

*__Corresponding author.__ Email: vsantorol@ub.edu

Abstract

Research projects involving science shops and citizen science in their promotion of participatory approaches are flourishing globally. However, an instrument evaluating the impacts of such approaches at different stages of a participatory research processes has yet to be validated. The InSPIRES H2020 project developed an impact evaluation tool for just this purpose, consisting of 64 items that reflect upon the dimensions of *knowledge democracy*, *citizen-led research*, *participatory dynamics*, *transformative change*, and *integrity*. In this article, we seek to test the content validity of this tool and to provide recommendations that can ensure its validity. A panel of nine experts was created to evaluate each item as regards the following three criteria: representativeness, relevance, and clarity. The Aiken's V and Wilson Score methods were used to assess the tool's content validity based on the experts' ratings. Experts' written comments were also reviewed. At the panel level, 75% of the items were considered satisfactory in relation to each of the three validity criteria. However, at the population level, 72% of the items suggested that parts of the tool were not valid and required revision. The main suggestions from the experts pointed to the need to reformulate items in which the separation between science and society appeared reinforced and to develop more items about the gender perspective of a research project. The revised version of the tool should serve as a well-founded, comprehensive evaluation instrument for on-going and future projects whose goal is to self-reflect and compare participatory research processes.

Key words: participatory research; impact evaluation; responsible research and innovation; citizen science; science shops

1. Introduction

In recent years, interest has grown in the application of participatory approaches, including citizen science (CS) and science shops (SS) in research projects (Kontic and Kontic 2018; Wiggins and Wilbanks 2019). These forms of public inclusion aim to make the community an essential agent in the definition of social concerns and the development of research proposals as well as in the performance of the scientific projects and the design of actions. As such, citizens and communities are no longer seen solely as objects of research, but rather as co-researchers involved in the production of knowledge

(Vohland et al. 2021). All in all, CS and SS as participatory approaches to research have led to a rethinking of the methods and frameworks of traditional research (Urias et al. 2020).

CS and SS can be considered similar concepts that aim to change the way science is produced by placing greater value on social needs and public engagement. SS, on the one hand, are intermediary units that seek to provide independent participatory research in response to concerns expressed by civil society actors (Leydesdorff and Ward 2005). Traditionally based within universities, these interfaces re-articulate social concerns into research questions together with social actors and undertake participatory research projects to address them (Le Crosnier, Neubauer and Storup 2013). SS act at the first stage of the research process and allow institutions to be more responsive to civic needs and for the research agenda to be defined by a bottom-up, demand-driven approach. SS have been characterized by their potential to democratize knowledge, placing all stakeholders, including non-experts, at the same level and considering all their knowledge relevant (Urias et al. 2020). CS, on the other hand, refers to research projects in which citizens or non-professional scientists participate in different phases of the study, particularly in the data collection and data analysis (Woolley et al. 2016). CS has been characterized by its emancipatory potential, strengthening the bonds of the community and enhancing the reflexivity and autonomy of the collective (Woolley et al. 2016) and by its role in support of social transformation through the development of new science–society–policy interrelations (Bela et al. 2016). The approach has gained popularity with the proliferation of technologies such as mobile applications and wireless sensors (Newman et al. 2012). Both SS and CS operate at the junction between science and society.

This interest for increasing public participation in research is being expressed being not only by different disciplines within the academic community but also, by the lay public, civic organizations, and other stakeholders, including granting agencies and policy-makers (Wiggins and Wilbanks 2019). The benefits for society and scientists of using CS or SS are multi-faceted and, include a greater research capacity, enhanced scientific literacy, enriched scientific knowledge, stronger community-building, and less public distrust in science (Den Broeder et al. 2018; Kieslinger et al. 2018). However, the traditional criteria used to measure research impact, often related to productivity performance (e.g. number of publications and total citations per year), are thus less suited for evaluating the characteristics of CS and SS.

Prior work has been undertaken defining criteria, indicators, and methods for impact evaluation in CS and SS (Schlierf and Meyer 2013; Kieslinger et al. 2018; Phillips et al. 2018; Schaefer et al 2021). Yet, online evaluation tools that are capable of demonstrating the value of CS and SS remain limited and untested as regards their content validity (Gresle et al. 2019). An additional drawback of these evaluations is that they often rely solely on scientific investigators' feedback and fail to include that of the other stakeholders involved in the project, which ultimately leads to a bias in the evaluation studies (Gresle et al. 2019). It is clear that evaluations of participatory research need to be framed from a multidimensional and multi-stage perspective where the process itself is worth evaluating (Schaefer et al. 2021). Thus, an impact evaluation tool tailored for SS and CS should reflect upon the implementation of the project and the way in which the participatory concepts have been incorporated, so that it can, in turn, become an opportunity for critical reflection and experiential learning (Gresle et al. 2019). One reason for the current dearth of standardized evaluation tools for SS and CS projects might be that the latter are very context-specific and tend to focus on a range of different objectives (Schaefer et al. 2021). Nevertheless, a tool that can assess several dimensions of SS and CS remains of some relevance for helping projects evaluate those dimensions that might have been overlooked and for comparing their results with evaluations made of other projects. Moreover, such a tool can help overcome some of the barriers to conducting evaluations in SS and CS projects, most notably the lack of financial and human resources (Gresle et al. 2019).

This lack of validated CS and SS research impact evaluation tools is one of the issues tackled by InSPIRES, an H2020 project financed by the European Commission. In addition to this goal, InSPIRES works for the co-creation of research questions responding directly to social challenges in the health and environmental sectors, with a specific focus on gender parity and support for vulnerable groups (InSPIRES Project 2020). InSPIRES has developed an online evaluation tool to gauge the impact of research projects and process that incorporates components of SS and CS. The tool is characterized by its multidimensional conceptualization and categorization of the impact of participatory research underpinned by the principles of SS and CS (Milat, Bauman and Redman 2015; Gresle et al. 2019). The tool evaluates both the immediate and mid-term impacts of research, stressing real-world benefits and capturing the quality of the participatory process (Gresle et al. 2019). In the particular case of CS, which is a constantly evolving field, the tool seeks to gauge the degree of citizen involvement in a project, but it is flexible enough so as not to target a specific type of CS.

The main objective of this study was to determine the content validity of the InSPIRES impact evaluation tool. We analyzed how far each item of the tool could be considered clear, representative of and, relevant to and the dimension being measured and we also, evaluated the comprehensiveness of the tool. On the basis of this analysis, we were able to adjust the content of the tool, provide recommendations for a validated version and further the discussion on what should be evaluated as key concepts to reflect public engagement with research and research engagement with society.

2. Methods

2.1 Inspires impact evaluation tool: structure, dimensions, and sub-dimensions

The original tool contains a total of 16 questionnaires designed to capture the multiple viewpoints of four distinct respondent profiles: civil society members, project managers or SS/CS coordinators, students, and researchers (scientists). Each profile of respondents has to answer a questionnaire in each phase of the research: Beginning, Mid-term, End, and Six months after completion of the project. Each of the 16 questionnaires contains four to six quantitative items and at least one open-ended question. For the quantitative items, respondents have to answer on a seven-point Likert scale ranging from 1 'I completely disagree' to 7 for 'I completely agree'. Two items require respondents to identify the stages in which the community was actively involved and the stages they consider community participation to have been most effective. In total, there are 64 items (55 quantitative items and nine open-ended questions). A more extensive description of the tool can be found in Gresle et al. (2019).

All the quantitative items are classified under five dimensions and sub-dimensions. The first dimension, *Knowledge democracy*, refers to the production of knowledge from the perspective that

multiple epistemologies exist (Hall and Tandon 2015). *Knowledge democracy* includes the sub-dimensions of transdisciplinarity, data and results openness, and the project's scientific relevance. The second dimension, *Citizen-led research*, refers to a scientific research process driven by non-professional scientists. Sub-dimensions included in this dimension are the alignment of the project goals with community demands and the efficacy of community engagement in the scientific process. The third dimension, *Participatory dynamics*, refers to engagement practices and reflects the state of stakeholders' participation in these practices. Sub-dimensions of *Participatory dynamics* include the degree of community involvement and the quality and impact of engagement practices. The fourth dimension, *Transformative change*, refers to the extent to which fundamental shifts are achieved in the course of the project, leading to positive outcomes. As such, sub-dimensions include individual learning and personal growth, project sustainability, the societal responsiveness of the project and its impact on policies, programs, and/or services. The fifth dimension, *Integrity*, is concerned with the extent to which a project adheres to the moral and ethical values deemed essential for responsible research practices of research. It encompasses expectations alignment, the inclusion of vulnerable groups, the incorporation of ethical and gender perspectives, the transparency of data management, and the quality of available resources. More detailed definitions of each dimension and sub-dimension can be found in Supplementary Table S1. The formulation of the items, their associated dimensions and sub-dimensions, the particular phase in which they are evaluated as well as the respondent profile to which the items are put are described in Supplementary Table S2.

2.2 Study design

In this study, content validation of the tool was based on an expert assessment (Rubio et al. 2003). A panel of nine experts was selected by convenience and in accordance with their professional experience in the dimensions being assessed, including SS, CS, and research impact assessment. The mean number of years of expertise was 23 with a range from 4 to 40 years. The experts were from Western Europe, North America, and South America. More details about the characteristics of the panel can be found in Supplementary Table S3. The experts had to respond to an online content validation form to assess the tool's validity. The form was tailored to evaluate the representativeness, relevance, and clarity of each item as well as the comprehensiveness of the tool (Grant and Davis 1997). A total of 55 items were evaluated for these three criteria while nine open-ended questions were evaluated solely for relevance and clarity, since these questions were not representing one specific dimension. Experts were invited to rate each item on a four-point Likert scale. For instance, to assess the relevance of the item '*The project objectives meet the community demands*' in the dimension of *Citizen-led research*, 1 was connoted as an irrelevant item and 4 was connoted as a relevant item. Experts could also provide written comments clarifying their ratings and share additional insights to improve the tool's content. The validation form was distributed between 1 March and 30 April 2020. Experts received an email with a cover letter explaining the project and describing briefly the tool's content. Definitions of each dimension and sub-dimension were provided in the form to guide the experts. From the panel, all nine experts rated items 1 to 12 and eight rated items 13 to 64.

2.3 Data analysis

The quantitative analysis was based on the values generated from two calculations: Aiken's V and the Wilson Score method on asymmetric confidence intervals (CIs) (Wilson 1927; Aiken 1980). Aiken's V was calculated using the following mathematical formula: $V = X - l/k$, where V is the item validity value; X is the mean score of the ratings assigned by the experts; l is the lowest validity score of the rating scale (if 4-point Likert-scale, $l = 1$); k is the difference between the highest validity rating and the lowest validity rating that can be selected by experts (if 4-point Likert-scale, then $k = 4 - 1 = 3$).

The V value ranges from 0 to 1. The closer an item is to 1, the more that item is considered as being representative, relevant or clear. If $V = 1$, it means that all experts selected the highest possible rating, whereas if $V = 0$ it means that all experts selected the lowest possible rating (Penfield and Giacobbi 2004). In this study, a cutoff value of 0.75 was selected to retain an item. In total, three V values were calculated for each item, as three validity criteria (representativeness, relevance, and clarity) were being evaluated.

Furthermore, as V is influenced by sampling error, it is recommended that a range of possible values that it might assume be established (Dominguez-Lara 2016). However, V is rarely expected to be normally distributed; a denser distribution at the higher extremes of the rating scale is more likely to occur. Hence, we applied an asymmetric CI for each V (Merino and Livia 2009). In this study, the use of a higher Type I error rate ($\alpha = 0.10$) was preferred because of the relatively new content topic covered by this tool and the small sample size formed of the panel (Penfield and Giacobbi 2004). The asymmetric CI was calculated using the Wilson Score (Wilson 1927). This method is known to be accurate when there is a small sample size and is not dependent on a data normal distribution. The lower limit value (L) of the CI represented the threshold population value accepting the validity of an item in terms of its representativeness, relevance, or clarity. A lower limit value ≥ 0.70 indicated an acceptable content validity of the item. Items below this value were subject to modification. The software Visual Basic 6.0 using a SPSS syntax was used to calculate V values and the CI of 90% using the Score method (Penfield and Giacobbi 2004).

Additionally, each written comment made by the experts was retrieved and assigned to an item by the main researcher and, later, double-checked by another researcher from the InSPIRES working group. By way of follow-up, the experts were invited to make general comments about the tool or explain further their written comments on the validation form. Finally, once the validation content values had been calculated and interpreted, each item—in, particular those assigned a low value—were examined by the authors. Subsequently, the results were shared with three partners in the InSPIRES consortium (from the Europe, Latin America, and Africa). Each partner carried out separate reviews of the results with special reference to proposals for item elimination or modification. The resulting reviews were discussed by the second and third authors in order to bring together theoretical and practical reflections about the items that required major reformulation or elimination.

3. Results

3.1 The items' Aiken's V values and confidence intervals

At the panel level, the majority of items [75% (48/64)] recorded a $V \geq 0.75$ for all three validity criteria (see Table 1). At the panel level, 82% (45/55), 84% (54/64), and 94% (60/64) of the items

Table 1. Content validity of the items

Dimension		Representativeness		Relevance		Clarity	
	Item	V	L	V	L	V	L
Knowledge democracy	1	1	0.91	0.96	0.85	1	0.91
	2	1	0.91	0.96	0.85	0.96	0.85
	3	0.96	0.85	0.96	0.85	0.89	0.75
	4	0.89	0.75	0.89	0.75	0.81	0.66*
	5	0.85	0.71	0.74*	0.58*	0.78	0.62*
	6	0.89	0.75	0.85	0.71	0.78	0.62*
	7	0.89	0.75	0.78	0.62*	0.70*	0.55*
	8	0.70*	0.55*	0.74*	0.58*	0.85	0.71
	9	0.74*	0.58*	0.74*	0.58*	0.81	0.66*
	10	0.78	0.62*	0.85	0.71	0.89	0.75
	11	0.85	0.71	0.78*	0.62*	0.86	0.71
	12	0.70*	0.55*	0.70*	0.55*	0.74*	0.58*
Citizen-led research	13	0.88	0.73	0.88	0.73	0.88	0.73
	14	0.88	0.73	0.75	0.59*	0.75	0.59*
	15	0.83	0.68*	0.75	0.59*	0.79	0.63*
	16	0.83	0.68*	0.83	0.68*	0.83	0.68*
	17	0.79	0.63*	0.79	0.63*	0.83	0.68*
	18	0.75	0.59*	0.79	0.63*	0.83	0.68*
	19	0.58*	0.42*	0.54*	0.38*	0.83	0.68*
	20	0.67*	0.50*	0.71*	0.54*	0.75	0.59*
Participatory dynamics	21	0.79	0.63*	0.79	0.63*	0.79	0.63*
	22	0.79	0.63*	0.79	0.63*	0.79	0.63*
	23	0.88	0.73	0.88	0.73	1	0.90
	24	0.88	0.73	0.83	0.68*	1	0.90
	25	0.79	0.63*	0.88	0.73	0.96	0.83
	26	0.75	0.59*	0.75	0.59*	0.79	0.63*
	27	0.71*	0.54*	0.75	0.59*	0.88	0.73
	28	0.79	0.63*	0.79	0.63*	0.88	0.73
	29	0.88	0.73	0.75	0.59*	0.83	0.68*
	30	0.75	0.59*	0.75	0.59*	0.79	0.63*
Transformative change	31	0.92	0.78	0.75	0.59*	0.88	0.73
	32	0.92	0.59*	0.79	0.63*	0.88	0.73
	33	0.75	0.59*	0.75	0.59*	0.79	0.63*
	34	0.75	0.59*	0.67*	0.50*	0.75	0.59*
	35	0.79	0.63*	0.83	0.68*	0.92	0.78
	36	0.83	0.68*	0.83	0.68*	0.92	0.78
	37	0.83	0.68*	0.83	0.68*	0.96	0.84
	38	0.75	0.59*	0.75	0.59*	0.88	0.73
	39	0.96	0.84	0.96	0.84	0.96	0.84
	40	0.63*	0.46*	0.67*	0.50*	0.83	0.68*
	41	0.67*	0.50*	0.75	0.59*	0.96	0.84
	42	0.67*	0.50*	0.75	0.59*	0.96	0.84
Integrity	43	0.79*	0.63*	0.79	0.63*	0.88	0.73
	44	0.92	0.78	0.92	0.78	1	0.90
	45	0.92	0.78	0.92	0.78	0.96	0.84
	46	0.92	0.78	0.92	0.78	0.96	0.84
	47	0.83	0.66*	0.79	0.63*	1	0.90
	48	1	0.90	1	0.90	0.88	0.73
	49	0.96	0.84	1	0.90	0.88	0.73
	50	0.88	0.73	0.88	0.73	1	0.90
	51	0.88	0.73	0.83	0.68*	0.96	0.84
	52	0.92	0.78	0.83	0.68*	0.92	0.78
	53	0.79	0.63*	0.67*	0.50*	0.79	0.63*
	54	0.79	0.63*	0.79	0.63*	0.67*	0.50*
	55	0.79	0.63*	0.79	0.63*	0.67*	0.50*
Open-ended questions	56	–	–	0.88	0.73	0.88	0.73
	57	–	–	0.92	0.78	0.92	0.78
	58	–	–	0.88	0.73	0.75	0.59*
	59	–	–	0.92	0.78	0.92	0.78

(continued)

Table 1. Continued

Dimension		Representativeness		Relevance		Clarity	
	Item	V	L	V	L	V	L
	60	–	–	0.88	0.73	0.88	0.73
	61	–	–	0.79	0.63*	0.75	0.59*
	62	–	–	0.79	0.63*	0.79	0.63*
	63	–	–	0.88	0.73	0.88	0.73
	64			0.92	0.78	0.92	0.78

V: Aiken's values; L: lower limit value of 90% confidence interval. Items 1 to 12 were rated by all nine experts. Items 13 to 64 were rated by eight experts. When V is below 0.75 and or the L value is below 0.70, the value is indicated with an asterisk (*). Items 56 to 64 were not designed to evaluate a specific dimension and, thus, were not validated in terms of their representativeness.

recorded a satisfactory validity level in terms of their representativeness, relevance, and clarity, respectively.

However, when applying a 90% CI, a total of 63% (30/48) of the items presented an acceptable validation value ($V \geq 0.75$) at the panel level for all three (or just two in the case of open-ended questions) criteria. These items did not present a value of $L \geq 0.70$ for one or more of the three criteria (or two if qualitative items) (see Table 1). Moreover, when considering exclusively population-level values, 72% (46/64) of the items presented a potential lack of representativeness, relevancy, or clarity. And, when considering each criterion at the population level, only 42% (23/55), 36% (23/64), and 58% (37/64) of the items reached a satisfactory level ($L \geq 0.70$) of representativeness, relevancy, and clarity, respectively.

Overall, taking into account both V and L values, 28% (18/64) of the items did not require revision. Items requiring major revision were classified mainly in one of the following dimensions: *Knowledge democracy*, *Citizen-led research*, *Participatory dynamics*, and *Transformative change*. An extended results table showing the mean ratings and the upper limit values of the 90% CI can be found in Supplementary Table S4.

3.2 Written comments from experts

Several of the written comments referred to key aspects related to participatory research theory and practice (see Supplementary Table S5). For instance, two specific concepts—'transdisciplinarity' and 'scientific'—raised concerns amongst the panel. Experts expressed reservations with regard to the lack of representativeness and relevance of these terms as used in the dimension of *Knowledge democracy*. One expert felt that the term 'transdisciplinarity' could be misleading: '*It is not necessary for all projects to be transdisciplinary, so a low score on this can be put in perspective*'. According to the experts, the term 'scientific' might not be applicable to small-scale projects: '*I don't like the framing of this in terms of "scientific" at all. For us this would be very odd for the majority of our projects. The wording of Science [. . .] is very complicated [. . .], people will not understand this in the context of community-based projects. It feels to me that this is designed much more for the large scale projects rather than small scale*'. Similarly, one expert thought that this term might impede democratization of knowledge by retaining a term more closely related to the traditional institutional system: '*Those of us who live in the worlds of research and science have a kind of in-depth understanding of what these words mean. In a time when we are stumbling forward towards a knowledge democracy framework, we may need to learn to speak in new ways, even finding new ways to speak of 'science' [. . .]*'.

The experts also pointed out the need for the tool to explore and develop more items related to the sub-dimension of *Gender perspective*: '*What is meant by a "gender perspective"? Translate this question into more ordinary language. . .There may need to be many more about gender*'. They also suggested to re-articulate what the tool understands as 'Traditionally excluded groups': '*What are traditionally excluded groups? In my view, for science shops, it is important to include groups that would not have access to research in another way (e.g. because they have no money). But not all projects have to focus on minorities or less educated people?*'.

Experts also recommended questioning scientists—and not just civil society members—about the skills and knowledge they gain when undertaking research of this kind: '*An item should be added intended for scientists regarding knowledge and skills in engagement practices, for sure they learn something about how to communicate with civil society members, how to conduct focus groups and workshop etc*'. Likewise, the experts considered it important that the project managers should not be the only ones to be questioned about personal and financial research resources. All experts' comments are listed in Supplementary Table S6. Most comments seek to improve the clarity of items or propose that certain items be presented to other agents.

For most of the items identified as being non-satisfactory in terms of their validity values at the panel and/or population levels, sufficient information was included in the expert's comments to account for these low ratings. Additionally, among the 18 items with good validity both at the panel and at population levels, seven were subject to improvement given that they were specifically targeted by minor comments. This was the case of items 2, 3, 13, 39, 44, 46. In the case, for example, of item 3 '*The project research instruments are freely available to anyone that wants to access and re-use them*', it was suggested that speak of 'instruments' was irrelevant as the term might be understood to be limited to just by licenses, and thus, 'research methodologies' was suggested as being more adequate term in this instance.

4. Discussion

4.1 Recommendations to improve InSPIRES impact evaluation tool

Whereas several items reported an acceptable level of content validity in the opinion of to the panel (total of 75%), a number of items (72%) were identified as demonstrating a potential lack of representativeness, relevance, or clarity when evaluated at the population level. Moreover, it was noted that even for seven items with a satisfactory level of validity at both panel and population levels, experts gave additional suggestions for modification. Based on the results obtained, several recommendations were accepted for enhancing the

Table 2. Item revisions

Sub-dimension	Phase	Item #	Item (Respondent)	Revision status*	Item revised for validation (Respondent)
Dimension: Knowledge democracy					
Openness	3	1	The project results are made available to the stakeholders in an appropriate understandable format. (Project manager, Civil society member)	MA.	
	3	2	The project results are made available to the public at large in an appropriate understandable format. (Project manager, Civil society member)	MO.	Project results are made available to the general public in an appropriate understandable format. (Project manager, Civil society member)
	3	3	The project research instruments are freely available to anyone that wants to access and re-use them. (Project manager)	MO.	Research methodologies are described clearly, so others can adapt or re-use them. (Project manager)
	3	4	The project research databases are freely available to anyone that wants to access and re-use them. (Project manager)	MO.	Raw data that do not infringe privacy or other ethical constraints are made available in a fair way. (Project manager)
	3	5	The project research results are freely available to anyone that wants to access and re-use them. (Project manager)	EL.	
Knowledge integration	3	6	The project embraced a transdisciplinary perspective. (Project manager)	EL.	
	3	7	Several expertise had been involved in the project. (Project manager)	EL.	
	4	8	The project has contributed positively to improve the perception about the role of science in society. (All)	MO.	The project helped to better identify the ways research processes can be applied to respond to societal issues. (All)
	4	9	The project has contributed positively to improve the perception about the role of society in science. (All)	MO.	The project helped to better identify the ways societal knowledge and practices can be applied to improve research. (All)
	1-3	N1		AD.	The design and orientation of the project is not guided only by the ideas of academics or scientific experts. (All)
	3	N2		AD.	Discussions in the project were based on what was said, not who said it, and arguments were exchanged in a respectful and rational manner. (All)
Scientific relevance	3	10	The results of the project have contributed to generate relevant scientific knowledge. (Scientist)	MO.	The results of the project have contributed to generate new knowledge relevant for each stakeholder. (Scientist)
	4	11	The project's results actively contributed to the scientific discourse (via scientific publications, blogs, etc.). (Project manager, Scientist, Student)	MA.	
	2	12	The project is scientifically relevant. (Scientist)	MO.	The project will produce scientific results that are relevant and increase knowledge of the topic. (Scientist)
Dimension: Citizen-led research					
Community alignment	1	13	The project objectives meet the community demands. (All)	MO.	The project objectives meet community needs, concerns and priorities. (All)
	1	14	The participation of the community in the project can positively contribute to meet the community demands. (Project manager, Scientist, Student)	MO.	The participation of the community in the project can positively contribute to meet community needs, concerns and priorities. (All)
	2	15	Your participation in the project can positively contribute to meet the community demands. (Civil society member)	MO	Your participation in the project can positively contribute to meet community needs, concerns and priorities. (Civil society member)
	2	16	The project's objectives meet the community demands. (Civil society member)	MO.	The research objectives cannot be met without involving the community in the research process. (All)

(continued)

Table 2. Continued

Sub-dimension	Phase	Item #	Item (Respondent)	Revision status*	Item revised for validation (Respondent)
Dimension: Knowledge democracy					
Responsiveness to community engagement	1	17	The scientific objectives cannot be reached without involving the community in the scientific process. (All)	MO.	The research objectives cannot be met without involving the community in the research process. (All)
	2	18	You are able to contribute to the project by expressing a personal viewpoint. (Civil society member)	MA.	
	1	19	I have personal and/or professional experience in the Project topic. (Civil society member)	MO.	You are able to contribute to the project by expressing your knowledge and experience. (Civil society member)
	3	20	In which stages of the scientific process has it been most effective for you to engage the community? (All)	MA.	
Dimension: Participatory dynamics					
Degree of engagement	3	21	In which stages of the scientific process has the community been actively involved? (All)	MA.	
	3	22	You were involved or invited to the final communication activity where research results were presented. (Civil society member)	MO.	You were involved in or invited to participate in drawing the main conclusions of the project. (Civil society member)
Motivation	1	23	You are motivated to participate in the project. (All)	MA.	
	2	24	You are motivated to continue being involved in the project. (Student, Civil society member)	MO.	You are motivated to continue being involved in the project. (Scientists, Student, Civil society member)
	3	25	You are motivated to get involved in similar projects. (Scientist, Student, Civil society member)	MA.	
Satisfaction with the participation	3	26	The participatory dynamics, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, have been fluid and easy to carry out. (Project Manager, Scientist)	MO.	Participation in the workshops, group discussions, meetings, online/collaborative data collection process, has been easy and effective. (Project manager, Scientist)
	3	27	The participatory dynamics, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, have positively influenced the results of the research. (Project manager)	MO.	Participation in the workshops, group discussions, meetings, online/collaborative data collection process, has positively influenced the results of the research. (Project manager)
	2	28	You are satisfied with the participatory dynamics, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, of the project. (All)	MO.	You are satisfied with the participation activities, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, of the project. (All)
Impact of the participation	3	29	The participatory dynamics, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, have contributed to build more socially relevant knowledge. (Scientist)	MO.	The participatory activities such as workshops, group discussions, meetings, online/collaborative data collection process, have contributed to build more socially relevant knowledge. (Scientist, Civil society member)
	4	30	Thanks to the participatory dynamics, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, you have been able to address the problem. (Civil society member)	MO.	Thanks to the participatory activities, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, you have been able to address the problem. (Civil society member)

(continued)

Table 2. Continued

Dimension: Knowledge democracy					
Sub-dimension	Phase	Item #	Item (Respondent)	Revision status*	Item revised for validation (Respondent)
Dimension: Transformative change					
Self-improvement	1	31	You feel confident to contribute to the project. (Student, Civil society member)	MA.	
	3	32	You feel able to take ownership of the process. (Civil society member)	MO.	You feel confident to use the results to achieve your purposes. (Civil society member)
	4	33	Your participation in the project has changed your behavior when it comes to the project topic. (Civil society member)	MO.	Your participation in the project has changed your behavior and/or your attitude. (Civil society member)
	4	34	The participation in the project has fostered your confidence to contribute to science. (Civil society member)	MO.	Participation in the project has fostered your confidence to contribute to collaborative research projects driven by social needs. (Civil society member)
Knowledge and skills	2	35	You are learning new skills, knowledge and attitudes during the project. (Student)	MO.	You are learning or enhancing new skills, knowledge and attitudes during the project. (Student)
	3	36	You have learnt or enhanced new skills, knowledge and attitudes during the project. (Student)	MO.	You have learnt or enhanced new skills, knowledge and attitudes during the project (Phase 2 and 3). (Student)
	4	37	The skills, knowledge and attitudes acquired during the project have positively contributed to impulse your professional career. (Student)	MO.	The project has had an impact on your choice of further studies or job. (Student)
	4	38	Your participation in the project has increased your knowledge about the project topic. (Civil society member)	MO.	Your participation in the project has increased your knowledge about the project topic (Civil society member) Phase 3
	3	N3		AD.	You have learnt or enhanced new skills, knowledge and attitudes in engagement practices (e.g. how to communicate with civil society members, how to conduct focus groups and workshops, etc.). (Scientist)
Collective capacity	4	39	The project generated new research questions, new projects and/or proposals. (Project manager, Scientist, Student)	MO.	The project generated new research questions, new projects and/or proposals (Project Manager, scientist, student, civil society)
	4	40	Project like this one can increase the probability to get funding. (Project manager, Scientist)	MO.	Projects like this one can increase the probability of getting funding (Project manager, scientist, Civil society member)
Tangible impact	4	41	The publication of the results caused alternative policy, programme, process, product or service options to be considered. (Project manager, scientist, society member)	MO.	The dissemination activities and outputs of the research findings caused alternative policy, programme, process, product or service options to be considered. (Project manager, scientist, Civil society member)
	4	42	The publication of the results led to improvements in an existing policy, programme, process, product or service. (Project manager, scientist, Civil society member)	MO.	The dissemination activities and outputs of the research findings led to improvements in on existing policy, programme, process, product or service. (Project manager, scientist, Civil society member)
Dimension: Integrity					
Resource availability	1	43	The financial resources to conduct the project are available. (Project manager)	MA.	
	3	44	The financial resources were appropriate for the project. (Project manager)	MO.	The financial resources were appropriate for the project. (Project manager, Scientist)
	3	45	The personal resources were appropriate for the project. (Project manager)	MO.	The personal resources were appropriate for the project. (Project manager, Scientist)

(continued)

Table 2. Continued

Dimension: Knowledge democracy					
Sub-dimension	Phase	Item #	Item (Respondent)	Revision status*	Item revised for validation (Respondent)
Transparency	3	46	The community knows what the data is going to be used for, where the data is going to be stored and shared. (Project manager)	MO.	You have informed civil society members what the data is going to be used for, and where the data is going to be stored and shared. (Project manager)
	2	47	You are clearly informed about what the data is going to be used for, where the data is going to be stored and shared. (Civil society member)	MA.	
Gender perspective	1	48	The project promotes a gender perspective in the research process and results. (Project manager)	MA.	
	3	49	A gender perspective has contributed to improve the project results. (Project manager, Scientist)	MA.	
	2	N4		AD.	The project makes an effort to involve diverse gender experiences and give equal importance to each of them. (Project manager, Scientist and Civil society member)
	3	N5		AD.	The possible effects of the research in gender inequalities (gender role, access, control of resources, equality) were considered. (Project manager, Scientist and Civil society member)
Expectation alignment	1	50	Expectations are clearly defined and communicated at the beginning of the project. (Project manager, Civil society member)	MA.	
	3	51	The project met your expectations. (All)	MA.	
Inclusivity	2	52	The project includes perspectives and feedbacks from the community throughout the scientific process. (Project manager, Scientist, Student)	MO.	The project includes perspectives and feedback from the community throughout the scientific process. (Project manager, Scientist, Student, Civil society member)
	2	53	The project includes traditionally excluded. (Project manager)	MO.	The project includes traditionally excluded groups, groups that would not otherwise have access to research and/or particular groups expressing specific interests and needs. (Project manager)
Self-reflexivity	1	54	Analyzing your project design with Self-Reflection Questionnaire One (downloadable PDF document) has helped you to raise or ratify awareness on crucial decisions at this stage of the process. (Project manager)	MA.	
	3	55	Analyzing your project's closure with Self-Reflection Questionnaire Two (downloadable PDF document) has helped you to raise or ratify awareness on crucial decisions for future processes. (Project manager)	MA.	
Open-ended questions					
		56	What motivated you to participate in this project? Feel free to include any other comment about your impressions at this phase of the project. (All)	MA.	
		57	Is anything in the project developing differently than you expected? If so, what and why? What could be improved? Feel free to include any other comment about your impressions at this phase of the project. (All)	MA.	

(continued)

Table 2. Continued

Dimension: Knowledge democracy					
Sub-dimension	Phase	Item #	Item (Respondent)	Revision status*	Item revised for validation (Respondent)
		58	How would you describe the main outcomes of this project from your perspective and what has ensured the achievement of those outcomes? Without the organization that manage this project, would you have considered any other alternative to address your demand? if yes, which alternative would you have considered in order to address your demand? In the case of this different alternative, would you have expected the results to be different? Feel free to include any other comment about your impressions at this phase of the project. (Civil society member)	MO.	How would you describe the main outcomes of this project from your perspective and what has ensured the achievement of those outcomes? (Civil society member)
		59	How would you describe the main outcomes of this project from your perspective and what has ensured these achievements? Feel free to include any other comment about your impressions at this phase of the project. (Student, Scientist)	MA.	
		60	Now that the project has finished, how do you think its design and implementation could have been improved, as for example regarding the involvement of different stakeholders, use of different methods, organization of different activities? (Project manager)	MA.	
		61	Were there changes in your context that you could attribute to this project? If so, please describe them. Feel free to include any other comment about your impressions at this phase of the project. (Civil society member)	MA.	
		62	Please describe the main competencies that you acquired through this project. Feel free to include any other comment about your impressions at this phase of the project. (Student)	MA.	
		63	How would you describe the benefits of this project from your perspective and what has ensured these achievements? Feel free to include any other comment about your impressions at this phase of the project. (Project manager)	MA.	
		64	Were there changes in your teaching approach and/or research lines that you could attribute to this project? If so, please describe them. Feel free to include any other comment about your impressions at this phase of the project. (Scientist)	MA.	
		N6		AD.	What other types of knowledge products were produced? (Scientist)

*AD., Added; MA., Maintained; MO., Modified; EL., Eliminated.

validity of the tool as an evaluation instrument applicable to a range of different participatory research projects. The proposed revisions are listed and detailed using a before-and-after format in Table 2. In total, six items were added, 24 remained the same, 37 were modified, and three were eliminated. The following section examines the main changes introduced and the reasons for making them.

4.2 Items added

Based on the ratings and comments from the panel, one item needs to be added in to the sub-dimension of *Knowledge integration*. Experts argued that the term 'transdisciplinarity' was not relevant. Rather, what they considered important was an evaluation of whether a project incorporated a research process that combined different types of knowledge. This same point has been made by other scholars that research should promote a process in which mental models (i.e. personal thought processes based on individual beliefs and experiences) are shared (Körner et al 2016; Evans et al. 2019). In this way, it is possible to avoid one single area of knowledge being overrepresented in the research process. Thus, it was decided to add the following item for all four agents in phase 3: '*Discussions in the project were based on what was said, not who said it, and arguments were exchanged in a respectful and rational manner*'.

According to the experts' evaluation, another item needs to be added to the sub-dimension *Degree of engagement*. The items in this sub-dimension were designed to measure an individual's level of involvement in a project. However, as participatory projects can usually be considered as incorporating a continuum of engagement levels (i.e. in terms of the roles enacted and the direct involvement in research-related activities) (Spears Johnson, Kraemer Diaz and Arcury 2016), it is not relevant to determine whether a project includes a greater or smaller number of stakeholders but, rather, whether this level of engagement responds to the stakeholders' needs and expectations (e.g. the amount of time they are willing to invest and their level of commitment) (Bird et al. 2020). Indeed, Wiggins and Wilbanks (2019) reported that both contributory projects (in which citizens participate in a narrow range of research phase activities) and co-created projects (in which citizens are involved in the whole research process and in the decision-making) have their own benefits for citizens. Consequently, the tool lacked an item that could gauge whether people were satisfied with the level of stakeholder engagement. Thus, it was decided to modify an item put to all four agents in phase 2: '*You are satisfied with the participation activities, such as for example participatory workshops, group discussions, meetings, online/collaborative data collection process, of the project*'.

In line with the experts' comments, it was decided to include an item in the sub-dimension of *Knowledge and Skills* in the broader dimensions *Transformative change*. The experts were concerned that by associating co-learning solely with the civil society members, the tool risked increasing the gap between science and society. Therefore, the experts recommended evaluating the transformative changes undergone by the scientists too. A frequently identified limitation of academic researchers is that they lack skills to facilitate participatory research that address both socially and politically sensitive issues (Cargo and Mercer 2008). It was therefore decided to include the following item for scientists in phase 3: '*You have learnt or enhanced new skills, knowledge and attitudes in engagement practices (e.g. how to communicate with civil society members, how to conduct focus groups and workshops, etc.)*'.

According to the experts, the sub-dimension *Gender perspective* was not sufficiently developed to be representative; moreover, they found the term to be unclear. Indeed, a number of scholars report the need to include a number of questions to evaluate whether agender perspective and gender equality have been promoted throughout the research processes (Korskvik and Rustad 2020; Biglia and Vergés-Bosch 2016). Based on the evaluation questions proposed in the literature, two items were added for project managers, scientists, and civil society members, one in phase 2: '*The project makes an effort to involve diverse gender experiences and give equal importance to each of them*', and, one in phase 3: '*The possible effects of the research in gender inequalities (gender role, access, control of resources, equality) were considered*'.

4.3 Items modified

It was suggested that a total of 37 items be modified. Those requiring notable modifications are described below. As reported in the results, the inclusion of the terms of 'transdisciplinarity' and 'multiple expertise' (items 6 and 7) were queried by the panel as they tend to presuppose that it is mandatory for a participatory research to cover a wide range of professional disciplines. Thus, items 6 '*The project embraced a transdisciplinary perspective*' and 7 '*Several expertise had been involved in the project*' were combined and expressed as one sole item: '*The design and orientation of the project is not guided only by the ideas of academics or scientific experts*'. The suggested modification sought to capture the idea that not one single point of view was dominant in the research project. Indeed, a project may not be characterized by pluralistic expertise but rather on a research process that embraces both scientific and non-scientific contributions to knowledge creation (Nowotny 2003).

The idea of 'scientific relevance' was also highly criticized by the experts as not being representative of what the tool should measure. This concern was most notable in relation to item 10: '*The results of the project have contributed to generate relevant scientific knowledge*'. The term 'scientific' is open to misinterpretation and, in the eyes of the panel, appeared to be dismissive of the potential social value of the research (Dobrow et al 2017). Indeed, participatory research is both inherently scientifically and societally relevant (Cargo and Mercer 2008). Thus, relevance is better understood in terms of whether the scientific research is formulating and investigating the right questions, that is, questions that seek to generate findings which can elucidate new knowledge and opportunities for action or change which matter for all the stakeholders involved (Balazs and Morello-Frosch 2013). Item 10 was therefore modified as follows: '*The results of the project have contributed to generate new knowledge relevant for each stakeholder*'.

To respond to the concerns expressed by the experts regarding the evaluation of scientific knowledge and production implied in certain items, the following modification was suggested. According to the panel, items 41 '*The publication of the results caused alternative policy, programme, process, product or service options to be considered*' and 42 '*The publication of the results led to improvements in an existing policy, programme, . . .*' attached too much importance to the publication of results as final knowledge products. Indeed, both items were rated as being not representative. As discussed above, in the case of research applying in SS or CS,

alternatives to traditional academic realizations and bibliometric indicators are preferred. Similarly, such beliefs correspond with the principles set up by the San Francisco Declaration on Research Assessment (2020), which calls for a more flexible, quality-over-quantity-oriented approach to assessing research. Indeed, there is a keen interest among research funders and experts to investigate innovative ways of disseminating research outputs (Ross-Hellauer et al. 2018). Thus, both items were similarly modified to encompass all possible dissemination activities: '*The dissemination activities and outputs of the research findings caused alternative…*'. This would appear to be a fundamental issue for participatory research projects as they are more likely to value other forms of knowledge products generated by the research—examples of non-traditional outputs include factsheets, video clips, and social media coverage (Rau, Goggins and Fahy 2018)—than the more conventional ones, that is, such as academic publications, conference proceedings or reports (Mårtensson et al. 2016).

In the sub-dimension *Inclusivity*, the experts reported that while item 53 '*The project includes traditionally excluded groups*' was representative, it was not especially relevant as not all collaborative projects focus their attention on minorities. One expert claimed that *Inclusivity* should be interpreted as including groups that would not otherwise have access to research if they could not participate in collaborative projects. According to Sauermann et al. (2020), in CS projects, certain matters of concern could reflect the interests and needs of particular groups of citizens, but this does not mean that this group is marginalized or traditionally excluded (without this diminishing the relevance of the research). Thus, it was recommended that item 53 be modified as follows: '*The project includes traditionally excluded groups, groups that would not otherwise have access to research and/or particular groups expressing specific interests and needs*'.

4.4 Items eliminated

Based on the results, it was suggested that a total of four items be eliminated: items 5 '*The project research results are freely available to anyone that wants to access and re-use them*', 6 '*The project embraced a transdisciplinary perspective*', and 7 '*Several expertise had been involved in the project*'. These items failed to reach satisfactory levels of validity as regards their representativeness and/or relevance, while the partners of the InSPIRES consortium did not feel they merited being retained. Item 5 was eliminated on the grounds that it was too similar to items 1 and 2. Items 6 and 7 were eliminated and replaced by one single item as described in Section 4.3.

4.5 Strengths and limitations

The application of CI in the interpretation of the results—an approach that is not frequent in content validation studies—helped to produce more accurate interpretations of the validity scores. Additionally, as the panel comprised experts from different cultural backgrounds and areas of expertise, we were able to obtain different perspectives and interpretations. We also employed a stepwise methodology to avoid, as far as possible, any subjectivity as much as possible in the interpretations of the results. Despite the extensive literature examining validation strategies (Grant and Davis 1997; Rubio et al 2003; Polit, Beck and Owen 2007), published studies do not necessarily follow a homogeneous approach. Thus, a distinctive feature of the study published herein is that the methodology developed should prove useful for other projects seeking to validate evaluation instruments. In their review of different methods used estimating content validity, Almanasreh, Moles and Chen (2019) identified the benefits of using of a three-stage assessment a opposed to a single-stage process based on just one sole value. In addition, integrating open-ended questions and inviting the experts to suggest other items for the tool helped identify items that might otherwise have been omitted from the dimensions evaluated. The concepts discussed in this study should, therefore, also be informative for researchers active in the evaluation of participatory research.

However, it should be noted that this study is subject to a number of limitations. The variability of ratings between different groups of experts could not be considered here as our sample size was small. Thus, this study assumed no interaction between potential groups of experts based on their cultural–professional characteristics such as work language, field of expertise, and years of experience. Additionally, when an expert gave a rating that was very different from those awarded by the rest of the panel, this sometimes caused considerable dispersion in the ratings and, as such, the values did not always represent a concerted consensus (Merino-Soto 2018). The elimination of extreme scores would have probably resulted in an increase of several L values of the CI. This may explain why several items with satisfactory V values presented a non-satisfactory L value when applying a 90% CI. Moreover, although background information about the tool was shared with the experts, we could not ensure that all experts were fully familiar with the tool. Furthermore, a four-point rating scale was preferred over a five-point scale so as not to include ambivalent middle ratings but using such a scale may have resulted forced ratings, i.e. ratings that do not correspond with the experts' opinions, since no option was available for them to be neutral. Finally, this study was conducted to make the InSPIRES impact evaluation tool a tool of reference for other researchers. However, participatory projects are typically context-specific; they have different local dynamics, and so, approaches are likely to vary (Bergold and Thomas 2012). This requires that specific content be tailored for evaluation tool. Here, the danger is that the content validation process might have led to an oversimplification of the items in an effort to generalize the scope of the evaluation and ensure its adoption by a range of different participatory research projects. One way in which the tool might be made more personalized without decreasing its validity would be to enable project members to give specific weightings to the different sub-dimensions and items according to their project. In this way, the results could be more representative of to their contexts and priorities.

5. Conclusions

The participatory research paradigm is being increasingly embraced by many research disciplines globally. However, to self-reflect on their participatory practices, researchers in SS and CS need a valid instrument with which to evaluate the research process and its potential impact. However, such tools remain limited and, to date, no tests have been conducted to ensure their validity. This study is the first to validate the content of an online participatory research impact evaluation tool, namely the instrument developed by the InSPIRES project. The study has demonstrated that, at the panel level, most of the tools' content could be deemed valid in terms of three criteria: representativeness, relevance, and clarity. However, at the population level, the results indicated that a large number of

items required revision. The tool's items related to research integrity garnered more consensus as regards their relevance, while, items concerned with the democratization of knowledge were open to greater debate. This suggests that future evaluation research needs to investigate how different contexts might influence the understanding of CS and SS. Nevertheless, this study ensures that the inferences drawn from the impact evaluation tool are meaningful and serve to inform on-going and future research projects about various aspects of applying CS or SS. The online impact evaluation tool is currently being tested by some 20 real-life projects. In the stages of this research, a user-experience test and a reliability evaluation of the tool will be conducted. The outcomes of this evaluation can then be used by research projects to self-reflect and to undertake comparisons of their participatory research processes. Not only does this evaluation effort represent an opportunity to gauge the impact of SS and CS, but it should, to some extent, also provide a more coherent understanding of what it means to apply CS and SS in research projects.

Supplementary data

Supplementary data are available at *Research Evaluation Journal* online.

Acknowledgements

We acknowledge support from the Spanish Ministry of Science, Innovation and Universities through the 'Centro de Excelencia Severo Ochoa 2019-2023' Program (CEX2018-000806-S), and support from the Generalitat de Catalunya through the CERCA Program. M.J.P. research is supported by the Ministry of Health, Government of Catalonia (PERIS 2016-2010 SLT008/18/00132). Also, this study would not have been possible without the precious contributions of all the experts from the panel. The time they took to evaluate the content of the tool and to give the team the very enriching feedback indicate the great generosity among researchers in the field of participatory research.

Funding

This work was supported by the European Union's Horizon 2020 research and innovation programme [under grant agreement No 741677]. The European Commission is not responsible for any use that may be made of the information it contains.

Conflict of interest statement. None declared.

References

Aiken, L. R. (1980) 'Content Validity and Reliability of Single Items or Questionnaires', *Educational and Psychological Measurement*, 40: 955–9.

Almanasreh, E., Moles, R., and Chen, T. F. (2019) 'Evaluation of Methods Used for Estimating Content Validity', *Research in Social and Administrative Pharmacy*, 15: 214–21.

Balazs, C. L. and Morello-Frosch, R. (2013) 'The Three R's: How Community Based Participatory Research Strengthens the Rigor, Relevance and Reach of Science', *Environmental Justice*, 6: 9–16.

Bela, G., Peltola, T., Young, J. C., Balázs, B., Arpin, I., Pataki, G., Hauck, J., Kelemen, E., Kopperoinen, L., Van Herzele, A., Keune, H., Hecker, S., Suškevičs, M., Roy, H. E., Itkonen, P., Külvik, M., László, M., Basnou, C., Pino, J., and Bonn, A. (2016) 'Learning and the Transformative Potential of Citizen Science', *Conservation Biology*, 30: 990–9.

Bergold, J. and Thomas, S. (2012) 'Participatory Research Methods: A Methodological Approach in Motion', *Forum: Qualitative Social Research*, 13/1: Art. 30.

Biglia, B. and Vergés-Bosch, N. (2016) 'Cuestionando la Perspectiva de Género en la Investigación', *Revista d'Innovació i Recerca en Educació*, 9: 12–29.

Bird, M., Ouellette, C., Whitmore, C., Li, L., Nair, K., McGillion, M. H., Yost, J., Banfield, L., Campbell, E., and Carroll, S. L. (2020) 'Preparing for Patient Partnership: A Scoping Review of Patient Partner Engagement and Evaluation in Research', *Health Expectations*, 23: 523–17.

Cargo, M. and Mercer, S.L. (2018) 'The Value and Challenges of Participatory Research: Stregthening Its Practice', *Annual Reviews, Public Health*, 29: 325–50.

Den Broeder, L. Devilee J., Van Oers H., Schuit A.J. Wagemakers A.(..2018) 'Citizen Science for Public Health', *Health Promotion International*, 33/3: 505–14.

Dobrow, M. J., Miller, F. A., Frank, C., and Brown, A. D. (2017) 'Understanding Relevance of Health Research: Considerations in the Context of Research Impact Assessment', *Health Research Policy and Systems*, 15: 31.

Dominguez-Lara, S. A. (2016) 'Content Validity Using Aiken's V with Confidence Intervals: Contributions to Rodríguez et Al', *Archivos argentinos de pediatría*, 114: e221–3.

Evans, J. M., Palmer, K. S., Brown, A. D., Marani, H., Russell, K. K., Martin, D., and Ivers, N. M. (2019) 'Out of Sync: A Shared Mental Models Perspective on Policy Implementation in Healthcare', *Health Research Policy and Systems*, 17: 94–6.

Grant, J. S. and Davis, L. L. (1997) 'Selection and Use of Content Experts for Instrument Development', *Research in Nursing & Health*, 20: 269–74.

Gresle, A. S. , Cigarini A., de la Torre Avila L., Jimeno I., Bagnoli F., Dempere H., Ribera M., Puertas E., Perelló J., Pinazo M.J.. (2019) 'An Innovative Online Tool to Self-Evaluate and Compare Participatory Research Projects Labelled as Science Shops or Citizen Science'. In: El Yacoubi S., Bagnoli F., Pacini G. (eds.) *International Conference on Internet Science*, pp. 59–72. Cham: Springer.

Hall, B. and Tandon, R. (2015). *Are We Killing Knowledge Systems? Knowledge, Democracy, and Transformations*. <http://www.politicsofevidence.ca/349/> accessed 16 Apr 2021.

InSPIRES Project (2020) 'About InSPIRES'< http://inspiresproject.com/about-inspires/> accessed 3 Aug 2020.

Kieslinger, B., Schäfer, T., Heigl, F., Dörler, D., Richter, A., and Bonn, A. (2018). Evaluating Citizen Science: Towards an Open Framework. In: Hecker S., Haklay M., Bowser A., Makuch Z., Vogel J., & Bonn A. (eds.), *Citizen Science—Innovation in Open Science, Society and Policy*, pp. 81–98. London: UCL Press.

Kontic, B. and Kontic, D. (2018). *Baseline research and best practice report on participatory and community-based research* <https://www.scishops.eu/wp-content/uploads/2020/02/SciShops.eu_D2.1-Baseline-research-and-best-practice-report-on-participatory-and-community-based-research.pdf> accessed 3 Dec 2020.

Körner, M., Lippenberger, C., Becker, S., Reichler, L., Müller, C., Zimmermann, L., Rundel, M., and Baumeister, H. (2016) 'Knowledge Integration, Teamwork and Performance in Health Care', *Journal of Health Organization and Management*, 30: 227–43.

Korskvik, T. R. and Rustad, L. M. (2020) *What is the gender dimension in research? Case studies in interdisciplinary research* <http://kjonnsforskning.no/sites/default/files/what_is_the_gender_dimension_roggkorsvik_kilden_genderresearch.no_.pdf> accessed 3 Aug 2020.

Le Crosnier, H., Neubauer, C., and Storup, B. (2013) 'Sciences Participatives ou Ingénierie Sociale: Quand Amateurs et Chercheurs co-Produisent Les Savoirs', *Hermès*, n° 67: 68.

Leydesdorff, L. and Ward, J. (2005) 'Science Shops: A Kaleidoscope of Science–Society Collaborations in Europe', *Public Understanding of Science*, 14: 353–72.

Mårtensson, P., Fors, U., Wallin, S.-B., Zander, U., and Nilsson, G. H. (2016) 'Evaluating Research: A Multidisciplinary Approach to Assessing Research Practice and Quality', *Research Policy*, 45: 593–603.

Merino, C. and Livia, J. (2009) 'Intervalos de Confianza Asimétricos Para el Índice la Validez de Contenido: Un Programa Visual Basic Para la V de Aiken', *Annals of Psychology*, 25: 169–71.

Merino-Soto, C. A. (2018) 'Confidence Interval for Difference between Coefficients of Content Validity (Aiken's V): A SPSS Syntax', *Annals of Psychology*, 34: 587–90.

Milat, A. J., Bauman, A. E., and Redman, S. (2015) 'A Narrative Review of Research Impact Assessment Models and Methods', *Health Research Policy and Systems*, 13: 18.

Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., and Crowston, K. (2012) 'The Future of Citizen Science: Emerging Technologies and Shifting Paradigms', *Frontiers in Ecology and the Environment*, 10: 298–304.

Nowotny, H. (2003) 'Democratising Expertise and Socially Robust Knowledge', *Science and Public Policy*, 30: /:151–6.

Penfield, R. D. and Giacobbi, P. R. (2004) 'Applying a Score Confidence Interval to Aiken's Item Content-Relevance Index', *Measurement in Physical Education and Exercise Science*, 8: 213–25.

Phillips, T., Porticella, N., Constas, M., and Bonney, R. (2018) 'A Framework for Articulating and Measuring Individual Learning Outcomes from Participation in Citizen Science', *Citizen Science: Theory and Practice*, 3: 3.

Polit, D. F., Beck, C. T., and Owen, S. V. (2007) 'Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations', *Research in Nursing & Health*, 30: 459–67.

Rau, H., Goggins, G., and Fahy, F. (2018) 'From Invisibility to Impact: Recognising the Scientific and Societal Relevance of Interdisciplinary Sustainability Research', *Research Policy*, 47: 266–76.

Ross-Hellauer, T. et al. (2018) *Innovative Dissemination Methods: Good Practices and Lessons Learned: OPENing UP New Methods, Indicators and Tools for Peer Review, Impact Measurement and Dissemination of Research Results*, <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5beac2f88&appId=PPGMS> accessed 3 Aug 2020.

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., and Rauch, S. (2003) 'Objectifying Content Validity: Conducting a Content Validity Study in Social Work Research', *Social Work Research*, 27: 94–104.

Sauermann, H., Vohland, K., Antoniou, V., Balázs, B., Göbel, C., Karatzas, K., Mooney, P., Perelló, J., Ponti, M., Samson, R., and Winter, S. (2020) 'Citizen Science and Sustainability Transitions', *Research Policy*, 49: 103978–16.

Schaefer, T., Kieslinger, B., Brandt, M., van den Bogaert, V., (2021) Evaluation in Citizen Science: The Art of Tracing a Moving Target. In: Vohland K. et al. (eds.) *The Science of Citizen Science*. Cham: Springer.

Schlierf, K. and Meyer, M. (2013) 'Situating Knowledge Intermediation: Insights from Science Shops and Knowledge Brokers', *Science and Public Policy*, 40: 430–41.

Spears Johnson, C. R., Kraemer Diaz, A. E., and Arcury, T. A. (2016) 'Participation Levels in 25 Community-Based Participatory Research Projects', *Health Education Research*, 31: 577–86.

Urias, E., Vogels, F., Yalcin, S., Malagrida, R., Steinhaus, N., and Zweekhorst, M. (2020) 'A framework for Science Shop processes: Results of a modified Delphi study', *Futures*, 123.

Vohland, K., Land-zandstra, A., Ceccaroni, L., Lemmens, R., Perelló, J., Ponti, M., Samson, R., and Wagenknecht, K. (2021) 'The Science of Citizen Science. 1st ed. Springer International Publishing.Vohland

Wiggins, A., and Wilbanks, J. (2019) 'The Rise of Citizen Science in Health and Biomedical Research', *The American Journal of Bioethics*, 19/8: 3–14.

Wilson, E. B. (1927) 'Probable Inference, the Law of Succession, and Statistical Inference', *Journal of the American Statistical Association*, 22: 209–12.

Woolley, J. P., McGowan, M. L., Teare, H. J. A., Coathup, V., Fishman, J. R., Settersten, R. A., Sterckx, S., Kaye, J., and Juengst, E. T. (2016) 'Citizen Science or Scientific Citizenship? Disentangling the Uses of Public Engagement Rhetoric in National Research Initiatives', *BMC Medical Ethics*, 17: 33.