

MDPI

Systematic Review

Co-Production Performance Evaluation in Healthcare. A Systematic Review of Methods, Tools and Metrics

Marta Marsilio 1,*0, Floriana Fusco 10, Eleonora Gheduzzi 20 and Chiara Guglielmetti 1

- Department of Economics, Management and Quantitative Methods (DEMM), Università degli Studi di Milano, via Conservatorio, 7, 20122 Milan, Italy; floriana.fusco@unimi.it (F.F.); chiara.guglielmetti@unimi.it (C.G.)
- School of Management, Politecnico di Milano, via Lambruschini 4, 20156 Milan, Italy; eleonora.gheduzzi@polimi.it
- * Correspondence: marta.marsilio@unimi.it

Abstract: Co-produced practices and publications in the healthcare sector are gaining momentum, since they can be a useful tool in addressing the sustainability and resilience challenges of health systems. However, the investigation of positive and, mainly, negative outcomes is still confused and fragmented, and above all, a comprehensive knowledge of the metrics used to assess these outcomes is lacking. To fill this gap, this study aims to systematically review the extant literature to map the methods, tools and metrics used to empirically evaluate co-production in health services. The search took place in six databases: Scopus, Web of Science, Psych INFO, PubMed, Cochrane and CINAHL. A total of 2311 articles were screened and 203 articles were included in the analysis, according to PRISMA guidelines. Findings show that outcomes are mainly investigated through qualitative methods and from the lay actor or provider perspective. Moreover, the detailed categorisation of the quantitative measures found offers a multidimensional performance measurement system and highlights the impact areas where research is needed to develop and test new measures. Findings should also promote improvements in empirical data collection on the multiple faceted co-produced activities and spur the consciousness of the adoption of sustainable co-productive initiatives.

Keywords: co-production; co-creation; performance; evaluation; measure; outcome; performance evaluation system; healthcare

1. Introduction

Co-production is widely considered a promising tool for dealing with current challenges in the health sector [1–3], where resources are being significantly reduced. Conversely, patients' expectations of higher-quality services are growing amid increasing demand caused by an aging population and the rise in chronic diseases. This pressures healthcare systems, challenging their long-term sustainability. To this end, policy makers (e.g., [4]) have promoted the development of more patient-centred personalised care based on new relational models, in which patients, their informal caregivers and local communities share responsibilities with care providers, thus enabling them to feel part of a team and fostering the quality of services. Patients are asked to participate actively and act as consumer producers, next to and in interaction with healthcare professionals and other stakeholders in healthcare, such as health providers, general practitioners, social services etc. [5,6]. This approach calls for a "community-accountable health development system" that can coordinate stakeholders' interests [7]. These aspects also assume relevance for unexpected challenges like rare events and pandemics. In recent months, the need to revise existing healthcare and social care systems with broader perspectives considering the skills, competences and experiences of whole ecosystems has appeared even more crucial due to the COVID-19 outbreak and the consequent risks [8–10]. Co-production has been identified as one possible solution to manage COVID-19 and to deliver public services enhancing



Citation: Marsilio, M.; Fusco, F.; Gheduzzi, E.; Guglielmetti, C. Co-Production Performance Evaluation in Healthcare. A Systematic Review of Methods, Tools and Metrics. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3336. https://doi.org/10.3390/ijerph18073336

Academic Editor: Milena Vainieri

Received: 15 February 2021 Accepted: 19 March 2021 Published: 24 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

community resilience through self-helping neighbourhood and volunteering activities. The collaboration of citizens (such as self-quarantining or wearing masks) in the face of limited opportunities for enforcement or encouragement guaranteed the success of the measures adopted by the health system in dealing with increasing infections [11].

Involving different stakeholders and relying on their cooperation and resource integration, co-production accentuates the complex and adaptive nature of the health service system, posing some challenges but also extending opportunities [12,13]. One of the challenges that certainly arises is to evolve a performance management and evaluation systems capable of systematically considering diverse value perspectives that come into play [14]. In this regard, it must be emphasized that there is increasing interest in understanding the benefits of co-production for the healthcare ecosystem [15]. Extant knowledge shows that co-production positively impacts on specific single dimensions, e.g., provider costefficiency (e.g., [16,17]); health outcomes (e.g., [18]); perceived service quality (e.g., [19]); service accessibility (e.g., [20]); customer satisfaction and quality of life (e.g., [21,22]); and compliance (e.g., [23]). Nevertheless, a systematic evaluation of the impacts of coproduction on the different stakeholders involved and its sustainability over time still is lacking [24,25]. This could also be ascribed to the "magic" nature of the co-concepts and the normative assumptions behind them, according to which a co-produced service ought to be a "better service" [2,26]. Moreover, an outcomes evaluation could be performed with different methods, indicators and timeline. Some recent reviews show that often the co-production effects are limited to narrative case studies, studies with small sample sizes and with poor attention to their sustainability [25,27]. Consequently, a comprehensive and robust co-production performance evaluation system would support in identifying the most suitable research approaches, methods and metrics for each specific outcome. This would also help to question the "celebratory" nature of co-production and to advance the investigation of unsuccessful and negative cases by challenging the "enchanting" nature of co-concepts and their effectiveness [26,28].

This paper aims to fill this gap. Drawing from the recent systematic categorisations of co-production outcomes according to all actors involved (lay actors, regular service providers and communities) [29], it systematically reviews the current literature to identify the methods, tools and metrics used to evaluate health co-produced services, according to each stakeholder. The paper offers a blueprint multidimensional performance measurement system that factors in the values of the multiple stakeholders and highlights the impact areas where further research is needed to develop and test new measures. The findings could promote improvements in empirical data collection on multiple faceted co-produced activities and spur consciousness of the adoption of sustainable patient-based initiatives.

The structure of this article is organised as follows. First, the research strategy adopted to conduct the review is presented. Subsequently, the main findings are reported and discussed. Finally, the paper provides recommendations and guidelines for future studies on co-production in the public and healthcare sectors.

2. Materials and Methods

To perform a replicable and transparent systematic review analysis, the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) have been adopted for the paper selection phase [30,31]. A PRISMA statement is an evidence-based minimum set of items for reporting systematic reviews and meta-analyses, useful for ensuring the rigor of systematic searches and that all relevant literature is included to decrease selection bias. The eligibility criteria and the four-phase flow of selection process are detailed below.

2.1. Study Eligibility Criteria

The objective of this study is to assess which research methods, tools and metrics have been adopted in the extant literature to measure the effects of co-production in the healthcare and public sectors. PRISMA guidelines distinguish between study eligibility

criteria and report eligibility criteria. The first criterion concerns the characteristics of the study that allow its inclusion or exclusion.

- Topic: Included studies must be focused on evaluating the effects of co-production. Moving from Nabatchi et al.'s (2017) [32] co-production definition, in this study we include a wide variety of activities in which service providers (i.e., healthcare providers) and lay actors (i.e., patients and/or carers) voluntarily work together in any phase of the health service cycle (i.e., commissioning, design, delivery, and assessment). As the concepts of "co-production" or "co-creation" are strictly related [33,34] and given their essential similarity [35], the systematic review includes both the literature on co-production and on co-creation. Considering both concepts adds useful insights to the outcome evaluations.
- Aim of the research: Only empirical studies targeting co-production/co-creation outcomes are included. Therefore, conceptual papers, reviews and protocols are excluded. No limitations on study design were introduced.
- Field: Both the health and public sector co-production literature were investigated. The second stream has been considered well-developed [3], and recent reviews on co-production in public show that health services emerge as one of the main sectors of interest [29]. Moreover, co-production outcome evaluation in public services (with a particular emphasis on health) has recently gained the main attention of distinguished scholars in the field [15,36]. Hence, this multidisciplinary approach can provide valuable insights into health sector co-production, with the caveat of considering contextual distinctions between the two sectors.

2.2. Search Strategy and Report Eligibility Criteria

Data were extracted from the largest peer-reviewed research literature databases in both fields (Scopus, Web of Science, Psych INFO, PubMed, Cochrane and CINAHL) in November 2020. The query was shaped with a three-level structure: the first defines phenomena; the second, the context; and the third, the unit of analysis. The choice of keyword limits was accurate for ensuring the completeness of the research, but also for focusing on the topic [37]. As displayed in Table 1, two queries were launched, respectively, concerning the health or public sector context, searching keywords in the string "topic" or "abstract, title and keywords", according to the specific tag used by each database.

Levels	Keywords_Health Dataset	Keywords_Public Dataset
First level (AND)	Co-product* OR Co-creat* OR Coprod* OR Cocreat*	Co-product* OR Co-creat* OR Coprod* OR Cocreat*
Second level (AND)	Health*	"Public Servic*" OR "Public Sector*"
Third level	Evaluat* OR Impact* OR Assess* OR Outcome* OR Indicator* OR Measur*	Evaluat* OR Impact* OR Assess* OR Outcome* OR Indicator* OR Measur*

Table 1. Search query structure.

Moreover, the following report eligibility criteria were selected:

- Language: Only English records are included, since it is the most-used language in the scientific literature and due to practical difficulties of translation.
- *Type of studies*: The review considers only articles published in scientific and peer-reviewed journals to ensure the reliability of sources and results.
- Publication status: Both published articles and, if available, articles in press were included to embrace the largest number of publications existing at the date of the search.
- Year of publication: No filters are adopted on the years of publication.

This strategy allowed the identification of 4192 records in the healthcare context and 492 in the public context.

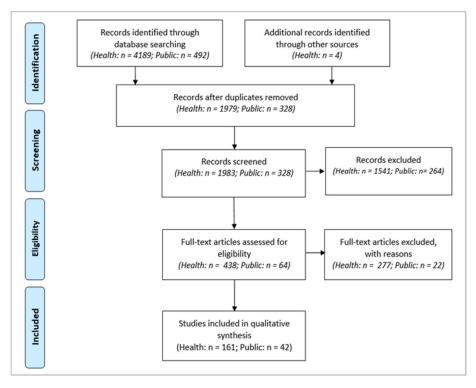


Figure 1 shows the PRISMA flowchart used to map the selection procedure.

Figure 1. PRISMA flowchart.

After record identification and duplicate removal, the screening phase was manually conducted on titles and abstracts. In the health dataset, 438 articles were considered eligible for full-text assessment, including 4 articles added with the snowball technique. In the public dataset, 64 articles were considered eligible for full-text assessment. In the second step of the screening phase, in the health dataset, 277 articles were excluded with reasons (e.g., no evidence for co-production/co-creation effects, no patient or caregiver inclusion in processes), and 161 articles were included in the qualitative analysis. In the public dataset, 42 were included in the qualitative analysis, while 22 were excluded. Among those 22 excluded, 8 focused on the health context and were already included in the healthcare dataset. In each of the two phases, first, all authors separately screened a sample of articles to verify the level of accordance in eligibility assessment; then, the second and third authors selected the included articles. All authors approved the final dataset, which consisted of 203 papers. Disagreements were resolved through discussion and consensus.

2.3. Data Collection

First, baseline information was extracted: author/s; title; journal; year; research questions; methodological approach (qualitative/quantitative/mixed); study design (e.g., case study, randomised controlled trial etc.); country of the study; field of the study (health/public) and context (e.g., mental health, primary care etc.).

For a detailed examination of the methods, tools and metrics used, articles were initially coded through Cepiku et al.'s (2020) outcomes identification and classification framework [29]. Specifically, within a comprehensive framework of co-production activation, management and evaluation for public services, they classified outcomes according to the actors affected by co-production: lay actors, service providers and the community at large. Considering the peculiarities of the health sector [27,38], a specific actor category was added to include outcomes for professionals, whose role and involvement in the co-production of the health service assume a specific role that must be distinctly analysed [27,39]. The focus on healthcare workers as a necessary part of the relationship between organisational processes and quality of care is not new in the literature [40].

Parkinson's (2018) systematic review stressed the importance of the health of workers and their families as a competitive advantage as healthcare organisations strive to deliver resource-efficient, high-quality care to patients [41].

Figure 2 represents the outcomes coding framework used to analyse the dataset.

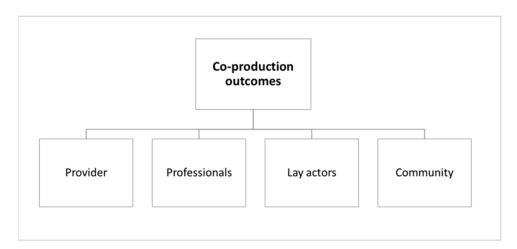


Figure 2. Outcomes' coding framework.

3. Results

3.1. Article Overview

The 203 articles included in the database cover the period from 1987 to November 2020. Figure 3 depicts the publication trends. The interest in the public field dates back to 1987, while in healthcare, the first article meeting the selection criteria is dated 1998. Attention to the co-production concept, both in the public and in healthcare fields, is growing exponentially: 69.5% of healthcare and 73.8% of public sector studies have been published in the last four years.

Considering journals, high fragmentation regarding articles per journal characterised both sub-datasets. Healthcare articles were published in 96 journals, of which only 27 (28.4%) have published at least two papers and only 18 have published at least three papers (11.2%). Research addressing the public sector has been published in 35 journals, of which only three (7.1%) had at least two articles (see Supplementary Materials Table S1).

Empirical evidence has been assessed in several countries, proving the adaptability and compatibility of these concepts within different contexts [3]. A strong geographical concentration is still detectable: 73% of health studies and 62% of public sector studies refer to five countries. This trend appears to be greater within the health field, in which almost 50% of the studies analysed are implemented in the UK (see Supplementary Materials Table S2). Considering the context of the research in the public domain, health and social care has the highest number of papers (11), followed by education (10) and general government (8), see Supplementary Materials Table S3. This result confirmed our methodological choice to include the public domain in the analysis as informative for healthcare investigation. Looking deeper into the healthcare domain, 61% of the papers deal with chronic or long-term disease patients. Another recurrent context is public health, especially concerning health prevention and promotion and the reduction of health inequalities in access to care. Increasing interest in the use of e-health and assistive technology is evident across all contexts (Table 2).

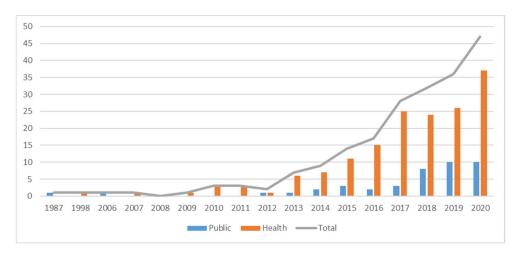


Figure 3. Annual scientific production—health and public.

Table 2. Studies' contexts in the healthcare domain.

	Context	n. Papers
Chronic or long-term conditions	Mental health, Elder care, Cancer, Diabetes, Heart disease, Rare chronic disease	100 (11 on e-health and 2 on assistive technology)
Public health	Health prevention and promotion, Health inequalities (cancer screening, maternity, mental health, healthy lifestyle, palliative care)	23 (2 on e-health)
Acute conditions	Osteoarthritis, Acute coronary syndrome, Self-harm	6 (1 on e-health)
Others	Integrated care, Primary care, Secondary care, Homecare in rural areas, Multiple settings	13 (2 on e-health)
Not specified		19 (3 on e-health)

3.2. Performance Evaluation System: Research Design, Approach Methods and Tools

The articles selected for review were analysed to assess which research methods, tools and measures have been used by authors to evaluate the impacts of co-production in the healthcare and the public sectors.

First, the research design and approach were considered. The adopted research design is clearly stated in a scant number of works. A full reading of the papers identified 34 with a longitudinal design, 8 with a cross-sectional design and 10 with a mixed design.

As depicted in Table 3, about half of the reviewed articles adopted a qualitative approach, while the other half was equally divided between quantitative and mixed approaches. There are some distinctions between the two domains regarding quantitative and qualitative percentages, but they may not be considered remarkable due to the varied sizes of datasets.

Research Approach	Total I	Dataset	Heal	thcare	Pu	blic
	n.	%	n.	%	n.	%
Qualitative	96	47.3	72	44.7	24	57.1
Quantitative	54	26.6	47	29.2	7	16.7
Mixed Methods	53	26.1	42	26.1	11	26.2
Total	203	100	161	100	42	100

Second, considering the number of stakeholders involved in the evaluation analysis, 47.8% of papers assessed the impacts only for one stakeholder, while 52.2% assessed the impacts for two or more stakeholders (Table 4). This evidence is common to both the health and public domains.

Table 4. Mono-stakeholder/multi-stakeholder evaluation approach.

	Mono-Stakeholder	Multi-Stakeholder
Health	47.8%	52.2%
Public	42.9%	57.1%

Third, to understand which are the main stakeholders considered in the co-production evaluation analysed and which are the methods of data analysis adopted, for each paper of the dataset, stakeholder outcomes (i.e., outcome for provider, outcome for professional, outcome for lay actors, outcomes for community) were identified and investigated. In total, 348 stakeholder outcomes occurred, 274 in the health domain and 74 in the public domain. On average, each paper reports impacts for 1.7 stakeholders. The results are summarized in Table 5.

Table 5. Methods for stakeholder outcome evaluation.

Methods of Data Analysis	_	utcome i Provide		_	utcome : ofession		_	utcome i		_	utcome i ommuni	
	Н%	P%	T%	Н%	P%	<i>T</i> %	Н%	P %	T%	Н%	P%	<i>T</i> %
Total	32	43	35	18	15	17	44	31	41	6	11	7
Quantitative	46	30	41	29	10	26	46	38	45	24	18	22
Qualitative	54	70	59	71	90	74	54	62	55	76	82	78

H% refers to health publications only; P% refers to public publications only; T% refers to the overall publications. In bold, the total outcome per stakeholder.

The most evaluated impacts concerned the lay actors (41%) and the provider (35%), while the impacts on professionals and on the community are largely under-investigated.

Considering the method of data analysis, there is a prevalent adoption of qualitative methods for all clusters of outcomes; however, some distinctions need to be highlighted. The evaluation of the outcomes for providers and lay actors is characterised by a high presence of quantitative analysis methods, whereas the outcomes for professionals and for a community have been assessed mainly through qualitative methods. Overall, the use of a quantitative approach is relatively more frequent in the healthcare domain.

Quantitative data analysis methods include statistical analyses, such as structural equation modelling (e.g., [42–44]), factor analyses (e.g., [44,45]), correlation and regression analyses (e.g., [6,16,46]), analysis of variance (ANOVA or MANOVA) (e.g., [47]), nonparametric test (e.g., [48–50]), multiple imputation techniques (e.g., [51]), and cost analysis, such as cost minimisation analysis (e.g., [16]) and cost opportunity analysis (e.g., [17]). Qualitative data analyses mainly embrace content analysis (e.g., [52]) and deductive or inductive thematic analysis (e.g., [53–55]), including specific inductive methodology, such

as Grounded Theory (e.g., [56–58]), Gioia methodology (e.g., [59,60]) and interpretative phenomenological analysis (e.g., [61]).

Finally, tools used in co-production evaluation were investigated. Quantitative data collection tools consist of surveys, questionnaires and secondary data, e.g., archival administrative data. Qualitative data collection tools included surveys, questionnaires, in-depth or semi-structured interviews, focus groups, meetings, workshops and observations. It should be emphasised that 16 papers adopted a narrative approach or unspecified quantitative and/or qualitative data collection and analysis methods.

3.3. Co-Production Outcomes: Quantitative Measures and Metrics

The focus was then narrowed to the quantitative measures and metrics used in quantitative or mixed papers for assessing the effects of co-production. More precisely, for each stakeholder (i.e., providers, professionals, lay actors, the community), diverse outcome analytical components were identified and the type of measures/metrics specified. Specifically, the paper moves from the analytical components of outcomes identified for each actor (e.g., for lay actor enjoyment and satisfaction, empowerment etc.; for provider efficiency, effectiveness etc.; for the community value for society etc.) as in the work of Cepiku et al.'s 2020 [29]. Considering the peculiarities of the health sector [38], some adaptations resulted in the analytical components' classification; Figure 4 shows the framework with the analytical components identified.

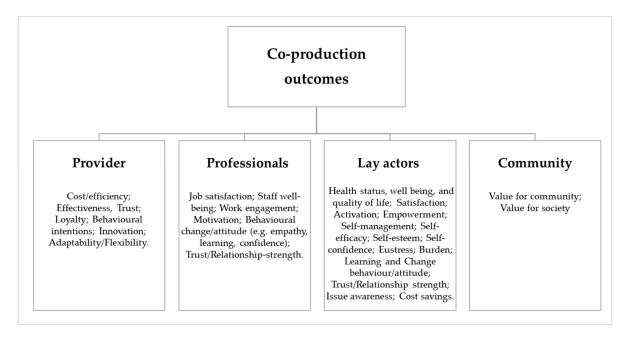


Figure 4. Analytical components framework.

The measure and metrics have been grouped according to the following categories: international validated scales, other tools (i.e., other reliable tools, whose robustness has been statistically tested in a paper or recognised in a specific field; adapted instruments from other works; and ad hoc instruments that include single or group of items created ad hoc for a study) and single indicators (i.e., specific measures adopted as a proxy to evaluate impacts).

Tables 6–9 summarise the results, proving analytical references for each component, measure or metric.

The following sections discuss the main results for the healthcare sector, highlighting the public literature when providing valuable additions.

3.3.1. Outcomes for Provider

Five main analytical components of outcomes for providers have been measured: (i) cost-efficiency; (ii) effectiveness; (ii) trust; (iv) loyalty; (v) behavioural intentions; (vi) innovation and (vii) adaptability/flexibility (Table 6).

Cost-efficiency (productivity) is considered the relationship between the costs of input and the related output [29]. This category includes cost saving and other efficiency outcomes. Cost savings have mainly been assessed through cost-benefit analysis. For instance, Span et al. (2018) developed a cost-minimisation analysis to calculate the total cost for hospital-based and home-based health strategies [16]. Medical (i.e., drugs and visits) and non-medical costs (transportation costs and work/school days missed) have been considered. Other studies used ad hoc indicators to calculate cost savings (e.g., [62]). Efficiency outcomes have been measured through indicators, such as the reductions for occupied bed days and admissions (e.g., [63]).

Effectiveness is the most assessed dimension for provider outcomes. It can be defined as the capacity to achieve the planned results in providing services. The effectiveness dimensions identified are service improvement, feasibility/acceptability of interventions or tools, and the usability of e-health tools. Service improvement has been evaluated mainly regarding increased accessibility or utilisation of a service through the use of specific indicators (e.g., [64,65]). This category assumes a particular meaning in healthcare, where the service is primarily effective when it generates an improvement or at least does not worsen the patient's health. However, for classification purposes, these types of outcomes (i.e., health status, quality of life and well-being) have been considered adopting the lay actors' perspective, given that they are the main beneficiaries of the service. In assessing service improvement, public sector findings go beyond the mere analysis of the increasing number of service users; they also assessed other features of the service, such as its frequency, distance from users [64] and performance (e.g., average examination course score) [66,67]. Several papers evaluated the usability of co-designing e-health tools through the System Usability Scale (SUS, [68]) (e.g., [69]); the Patient Education Materials Assessment Tool (PEMAT, [70]) has served to check the comprehensibility and actionability of health education material co-created by expert clinicians and patients in Badiu et al. (2017) [71]. Within effectiveness indicators, the percentage of patients declining (or accepting) an intervention (e.g., health screening) can be found [72].

Trust and *loyalty* are assessed mainly through ad hoc questionnaires, resulting from the previous literature (e.g., [73,74]).

Behavioural intentions, such as recommendation intentions and positive word of mouth, have also been evaluated through other existing and validated tools, such as the behavioural intentions construct of Dagger et al. (2007), for example [75].

Few measures have been found on *innovation*, measured by Sehgal and Gupta (2020) with scale items created ad hoc, according to the extant literature.

Adaptability and flexibility of services, investigated concerning the decentralisation of power and higher response to users' need in Dhirathiti (2019) [76] and McAllister et al. (2018) [77], have been measured, respectively, with ad hoc items and a validated scale. While these dimensions in the health sector have been mainly studied with micro and meso perspectives investigating changes at the patient–professional and provider levels, in the public sector, the analysis shifts at the government macro level by investigating the reorganisation of the governmental structure regarding the decentralisation of power [76].

Table 6. Measures and metrics of outcomes for provider.

Outo	come	International Validated Scales	Other Tools	Indicators	Main References
	Cost savings		Cost—benefits analysis (cost-minimization analysis [16]; opportunity-cost analysis) [17];	Length of admission x average costs per night [62];	[16,17,62];
Cost-efficiency	Efficiency			Reductions of: occupied bed days; hospital admission; days on a community treatment and community contacts [63];	[63];
Service improvement			Ad hoc [64,78]*;	 N. person using the service [76] *; N. persons screened/target [20,72]; Offering a specific type of service or not [65]; Students' examination results and game scores [66,67] *; % of recycling on the total amount of waste achieved by each municipality [79] *; Count variable calculating by summing the pro-environmental activities undertaken [80] *; 	[20], [64] *, [65], [66] *, [67] *, [72]; [76] *, [78] *, [79] *, [80] *;
Effectiveness	Usability of e-tools	System usability scale (SUS) [69,81,82];	Ad hoc [83], [84] *;	N. of user requests for service support [84] *;	[69,81–84] *;
	Acceptability/ feasibility of interventions/ e-tools	 Social Support Programme Acceptability Rating Scale [48]; Patient Education Materials Assessment Tool (PEMAT) [71]; Training Acceptability Rating Scale (TARS) [52]; 	Ad hoc [85,86];	N. or % of patients declining (or accepting) the intervention [72];	[48,52,71,72,85,86];
Trust			Ad hoc [73], [74] *;		[73], [74] *;
Loyalty			Ad hoc [73];		[73];
Behavioural intentions			 Construct of Dagger et al. (2007) [45,75]; Construct of Zeithaml et al. 1996 [21]; Ad hoc [66,78] *, [23,43,87] 		[21,23,43,45], [66] *, [75], [78] *, [87];
Innovation			Ad hoc [42];		[42];
Adaptability/ Flexibility		CSHCN Stand-Alone Questionnaire [77];	Adaptation from Voorberg et al. 2015 [76] *;		[76] *, [77];

The asterisk * indicates publications in the public domain.

3.3.2. Outcomes for Professionals

The main outcomes of the analytical components concern: (i) job satisfaction; (ii) staff well-being; (ii) work engagement; (iv) motivation; (v) behavioural change; and (vi) trust in professionals/relationship strength (Table 7).

Job satisfaction is a widely investigated construct in the academic literature, with many definitions and available operationalisations. However, it seems to be scarcely measured as a co-production outcome. Only few works addressed this issue, with confounding results. Three of four studies found that the co-creation of care increases the job satisfaction of mental health professionals [6,88,89], while Den Boer et al. (2017) did not find any correlation [46]. Two papers adopted the Measurement of Job Satisfaction (MJS, [90]) as one of the most reliable, valid and multidimensional measures of job satisfaction, while the other two use ad hoc tools.

Job well-being concerns the holistic perspective of workers' physical and emotional status concerning their work environment. Levels of well-being have been assessed in some studies through the Social Production Function Instrument (SPF-IL scale, [91]), a 15-item validated scale including some relevant dimensions of subjective well-being (universal goals, affection, behavioural confirmation, status, comfort and stimulation). Burnout is measured through the most-known tool, the Maslach Burnout Inventory (MBI, [92]). Finamore et al. (2020) applied the MBI to verify the effect of co-produced personality disorder training on staff [93]. The same scale was used in Farnese et al. (2020) to investigate the effects of informal co-production between professionals and caregivers, but no correlation was found [94].

Work engagement and motivation are intended as active involvement and the willingness to perform one's job. To verify the effect of co-delivery training on professionals, Hastings et al. (2018) used the Staff Positive Contributions Questionnaire (SPCQ, [95]) to evaluate the impact on staff motivation. The well-known Utrecht Work Engagement Scale (UWES, [96]) was used to measure work engagement in Ding et al. (2019) [6], which aimed at verifying the effect of patient participation in value co-creation with hospital nurses.

Behavioural change considers all changes in skills and personal behaviour in clinical routines, including ways of relating to patients. Within this category, the dimension most assessed has been staff empathy and attitudes towards patients. They have been assessed mainly in studies focusing on co-delivery training for mental health staff. Among the validated scales adopted, Staff Empathy for people with Challenging Behaviour Questionnaire (SECBQ, [97]) can be mentioned, as well as Borderline Personality Disorder-Cognitive/emotional Attitudes Inventory (BPD-CAI/FAI, [98]). Self-efficacy, defined as the belief in one's ability to succeed in specific situations or accomplish a task, has been measured through other reliable tools and adapted instruments.

Trust is considered related to this stakeholder, as the relationship between lay actors and professionals (e.g., dyadic relationship such as patient–doctor). Findings show that it has been assessed only with ad hoc or adapted instruments (e.g., [73,99]). For this outcome dimension, public metrics assessed are limited in number and do not add any valuable and innovative insights to the health findings.

No specific indicators have been found for this stakeholder outcome.

Table 7. Measures and metrics of outcomes for professionals.

Outcon	ne	International Validated Scales	Other Tools	Main References
Job satisfaction		Measurement of Job Satisfaction (MJS) [46,88];	Ad hoc [6,89];	[6,46,88,89];
Staff Well-being (including burnout)		 Social production function scale (SPF-IL) [46,88]; Maslach Burnout Inventory [93,94]; 		[46,88,93,94];
Work engagement		Utrecht Work Engagement Scale (UWES) [6];	Co-productive-taxpayer-as-supervisor (CTS) construct [100] *;	[6], [100] *;
Motivation		Staff Positive Contributions Questionnaire (short version) [97];		[97];
	Helping behaviour		Adapted from Schneider et al. 2005 [6];	[6];
Pakarianal dana	Empathy/ Attitudes	 Personality disorder-knowledge, Attitudes and Skills Questionnaire [47,93,101]; Similarities and empowerment attitude sub-scales from the Community Living Attitudes Scale [97]; Staff Empathy for people with Challenging Behaviour Questionnaire (SECBQ) [97]; The borderline personality disorder-cognitive/emotional attitudes inventory (BPD-CAI/FAI) [98]; 	Attitude towards self-harm in CYP of Crawford et al. (2003) [102];	[47,93,97,98,101, 102];
Behavioural change	Learning	 Personality Disorder-Knowledge, Attitudes and Skills Questionnaire [47,93,101]; Borderline personality disorder-cognitive/emotional attitudes inventory (BPD-CAI/FAI) [98]; 	Adapted from Crawford et al. (2003) [103];	[47,93,98,101,103];
	Self- efficacy		 Challenging Behaviour Self-Efficacy Scale (CBSE) [97]; Adapted version of the Self-efficacy Towards Helping Scale [103]; 	[97,103];
	Clinical behaviour intentions	Continuing Professional Development Reaction Questionnaire [103];		[103];
	Confidence		Ad hoc [103];	[103];
Relationship- strength			 Adapted from Hall et al., 2001 [99]; Adapted from Rajah et al. (2008) [73]; 	[73,99];

The asterisk * indicates publications in public domain.

3.3.3. Outcomes for Lay-Actor

Outcomes of co-production on lay actors are the most investigated with quantitative approaches. Fifteen analytical components have been identified: (i) health status; (ii) satisfaction; (ii) activation; (iv) empowerment; (v) self-management; (vi) self-efficacy; (vii) self-esteem; (vii) self-confidence; (ix) eustress; (x) burden; (xi) learning; (xii) changes in behaviour/attitude; (xiii) relationship strength; (xiv) issue awareness; and (xv) cost savings (Table 8).

Health status, including well-being and quality of life, understood as physical, mental and social health and well-being, have been found to be the most measured outcomes

in the healthcare sector. These outcomes have been mainly assessed with recognised, validated, self-reported scaled and clinical objective indicators, such as blood biochemical parameters [104]. The most recurrent metrics are the Patient Health Questionnaire (PHQ-9, [105]) (e.g., [18]); Health of the Nation Outcome Scales (HoNOS, [106]) (e.g., [63,107]); and Warwick–Edinburgh Mental Well-being Scale (WEMWBS, [108]) (e.g., [109]). Given the scarce investigation of informal caregiver perspectives, it is interesting to mention the study of Wood et al. (2010) [110] that used the Quality of Life Questionnaire for Family (QLQ-F; [111]) and treatment Group (QLQ-G, [111]) for assessing the perceived quality of family life after co-created training for people with addiction disorder and their caregivers.

Satisfaction includes all the metrics assessing the lay actor's subjective perception of co-produced service experiences, outcomes and processes. Satisfaction with a service (mostly derived from performance and service outcomes) and with the process (mainly referring to the enjoyment of collaboration) have been considered. These dimensions have been measured both with validated scales and ad hoc or adapted scales, whose reliability and validity are generally verified in studies. Within the first tools, the Social Support Programme Acceptability Rating Scale ([112]) was used by Brown et al. (2020) [18] both to evaluate the acceptability (outcomes for provider) and satisfaction of a participant in community co-produced intervention. Ad hoc or adapted instruments draw on some key previous studies, such as the perceived value scale developed in Sweeney and Soutar (2001) [21] or the service quality scale developed in Dagger et al. (2007) [113]. Hau (2018) [43] measured outcome (perceived) values and process (perceived) value in health co-creation with Hau and Thuy's (2012) scales [114]. In the public domain, Sanina et al. (2020) [66] used an ad hoc scale for assessing students' satisfaction with the co-produced activities in which they were involved. Within metrics used to assess informal caregivers' satisfaction, the "adjusted version of the caregivers' satisfaction with inpatient stroke care (C-SASC, [115]) 11-item scale" [116] and Family Experiences with Coordination of Care (FECC, [117]) Measure Set [77] have been used.

Activation defines one's ability (knowledge, skills and confidence) to take independent actions in one's life, for example to manage their care. In the health domain, this dimension is often measured with the well-known Patient Activation Measures (PAM, [118]), such as in Turner et al. (2015) [119], which used this scale to assess the effect of co-produced self-management training. The activation in the health field has been mostly interpreted as confidence in self-management [119,120], while in the public sector it has been assessed as the level of activeness regarding participation [121].

It should be emphasised that the PAM has also been used for assessing perceived *empowerment*, as in Jo and Nabatchi, 2019 [99]. It could be interpreted as a status (or) process by which people gain control over their lives. Empowerment dimension includes metrics that assess lay actors' capacities to become co-producers of health and well-being for their own [99] or for their loved ones [109,122]. The empowerment dimension in the public sector has also been measured with an adapted scale, such as the Self-Report Level of Participation Survey, which aims at assessing the perceived sense of involvement of lay actors during co-design workshops [123].

Self-management has been defined as the confidence of patients in self-care. It has been assessed through validated scales, such as the Health Education Impact Questionnaire (heiQ, [124]) (e.g., [119,120]) and Summary of Diabetes Self-Care Activities ([125] in [126]). Other psychological benefits regarding self-efficacy, self-esteem and self-confidence have rarely been assessed through metrics within the healthcare and public sectors. Fors and colleagues [50,127,128] evaluated in a randomised control trial the effects of person-centred and co-created care in acute disease using the General Self-Efficacy Scale (GSE scale, [129]); Wood et al. (2010) [110] used Rosenberg Self-Esteem Scale (RSE, [130]), while, in the public sector, Sanina et al. (2020) studied self-efficacy by looking at students' confidence in their professional skills [66].

The positive and negative effects of co-production on stress have been assessed regarding *eustress* and *burden* with validated or adapted instruments in the healthcare field,

such as the Burden Assessment Scale [131], which was used in Chiocchi et al. (2019) [109] to evaluate the effect of co-delivery psychoeducation programme on carer burden.

Both in the public and health academic literature, *learning and behaviour change* have been widely recognised as a possible effect of co-production. In quantitative terms, learning outcomes have been assessed regarding health literacy with validated (i.e., the newest Vital Sign UK (NVS-UK, [132]) ([18]), Diabetes Knowledge Test (DKT, [133]) in New et al. (2010) [126]) or ad hoc scales, and also with specific indicators, such as the learning scores of teaching tests [66,134]. Other behaviour/attitude changes include (improved) compliance, explained as the extent to which patients (or other lay actors) follow service provider's instructions, decisional conflict, that is, a state of uncertainty about the actions to be taken or opinions to be expressed, or a change in life routine (e.g., eating style). They have been assessed mainly in the health domain, with a validated scale or ad hoc adapted instruments (e.g., [81,94]). Some indicators are also adopted (e.g., [72,135]). Even though the change of behaviour or attitude has been mainly assessed in the health sector, the ad hoc scale suggested by Chen et al. (2015) deserves attention since it also investigates possible negative changes of behaviours in lay actors, such as unethical and rude behaviours [100].

The last three dimensions are few and measured largely with ad hoc instruments.

Relationship strength concerns a better relationship with professionals; the bidirectional value of this outcome means that it is also used for lay actors. It has been measured with an ad hoc or adapted instrument, as already seen in the sub-paragraph on professionals' outcomes.

Issues awareness is only investigated in Jo and Nabatchi (2019) by measuring people's perception of the importance of a given issue with an ad hoc item [99].

Cost saving is mainly discussed in relation to providers; however, some authors also highlight that there are cost effects for lay actors. For instance, in their cost minimisation analysis, Spanò et al. (2018) also consider lay actors' perspectives, which include missed work/school days and travel expenses for visits [16].

 Table 8. Measures and metrics of outcomes for lay actors.

Outcom	ne	International Validated Scales	Other Tools	Indicators	Main References
Health status, well-being and quality of life	Physical and mental health status	 General Health Questionnaire-12 [48]; Health of the Nation Outcome Scales HoNOS [63,107]; Generalized Anxiety Disorder Questionnaire (GAD-7) [18]; Patient Health Questionnaire (PHQ-9) [18]; Hospital Anxiety and Depression Scale (HADS) [119,120,136,137]; EORTC QLQ-C30 [137]; Functional Assessment of Chronic Illness Therapy—Fatigue FACIT-F [137]; Insomnia Severity Index (ISI) [137]; Saltin Grimby Physical Activity Level Scale (SGPALS) [50]; Strengths and Difficulties Questionnaire [86,138]; Post-Traumatic Stress Disorder (PCL-C); [137]; Mood andFeelings Questionnaire (MFQ) [86]; Revised Child Anxiety and Depression Scale (RCADS) [86]; Western Ontario and McMaster University Osteoarthritis Index (WOMAC) [139]; 	Ad hoc [140];	 Improved/deteriorated or unchanged clinical indicators (e.g., waist-to-height ratio; carotid artery reactivity (CAR); blood biochemical parameters; death) [49,104,136]; Re-admission/re-hospitalization for unscheduled reasons; occurrence of an adverse outcome [50,141]; 	[16,18,21,22,44,45,48– 51,63,86,104,107,109,110,116, 119,120,135–143];
	Quality of life and health-related quality of life	 Quality Quality-Adjusted Life-Years (QALYs) using HAE-BOIS-Europe survey [16]; EuroQol index (EQ 5D index) and the EuroQol Visual Analogue Scale (EQ VAS) [119,120]; Quality of Life Questionnaire for family (QLQ-F) and treatment group (QLQ-G) [110]; 	Fox Simple Quality-of-Life Scale [21];		
	Well-being	 15-item version of the Social Production Function Instrument for the Level of Well-being (SPF-ILs) (both patients and carers); [51,116,142,143]; Warwick-Edinburgh Mental Well-being Scale (WEMWBS); [48,109,135,136]; McGill Quality of Life Questionnaire (MQOL index) [44]; McGill quality of life questionnaire (MQOL index)-revised [22]; 	Perceived well-being construct of Sweeney et al. (2015) [144];		_

 Table 8. Cont.

Out	come	International Validated Scales	Other Tools	Indicators	Main References	
Satisfaction	With service/care	 Social Support Programme Acceptability Rating Scale [48]; Patient Assessment of Chronic Illness Care Short version (PACIC-S) [143]; Psychological Needs Satisfaction in Exercise Scale [49]; Family Experiences with Coordination of Care (FECC) Measure Set [77]; Diabetes Management and Evaluation Tool (DMET) [126]; 	 Perceived service quality and satisfaction construct of Dagger et al., 2007 [45,75]; Perceived quality of service questionnaires of Veltro et al. (2007) [107]; Service satisfaction scale of Oliver 2010 [21]; Outcome value scale of Hau and Thuy, 2012 [43]; Adjusted version of the caregivers' satisfaction with inpatient stroke care (C-SASC) 11-item scale [116]; Adjusted version of the Satisfaction with Stroke Care questionnaire (SASC) [51,116,142]; Adaptation from: Rajah et al., 2008 [73]; Aiken and Patrician (2000) [94]; Yu and Dean (2001) [145]; Wang et al. (2004) and Sweeney and (2001) [145]; Sweeney and Soutar (2001) and Mathwick et al. (2001) [146]; Sweeney and Soutar (2001) [75,147]; Suarez-Alvarez et al. (2020) [148]; Ad Hoc [64,66,74] *, [6]; 		[6,21,43,45,48,49,51], [64] *, [66] *, [73], [74] *, [75,77,94,107,116,126,142, 143,145,146,148]	
	With process (enjoyment)			 Process value of Hau and Thuy, 2012 [43]; Adaptation from: Wang et al. (2004) and Sweeney and Soutar (2001) [145,147]; Nelson and Byus (2002) [75]; Ad hoc [144]; 		[43,45,75,145,147]
Activation		Patient Activation Measure [119,120,135]	Ad hoc [121] *;		[119,120], [121] *, [135]	
Empowerment		 Family Empowerment Scale [77,109] Patient Activation Measure (PAM) [99]; SIS Communication/Behavior Skills Questionnaire [110]; Diabetes Empowerment Scale Short Form (DES-SF) [126]; 	 Adaptation from spidergram tool of Draper et al. (2010) [123] *; Ad hoc [122] *; 		[77,99,109,110], [122] *, [123] *, [126]	

 Table 8. Cont.

Outcome		International Validated Scales	Other Tools	Indicators	Main References
Self- management		 Health Education Impact Questionnaire (heiQ) [119,120]; Summary of Diabetes Self-Care Activities (DSCA) [126]; 			[119,120,126]
Self- efficacy		General Self-Efficacy Scale (GSE scale) [50];	Ad hoc [66] *;		[50], [66] *;
Self- esteem		Rosenberg Self-Esteem Scale [110];			[110];
Self- confidence			Ad hoc [149];		[149];
Eustress			Adaptation from: Eustress construct of Simmons and Nelson 2001/Nelson and Simmons 2003 [23];		[23];
Burden		Burden Assessment Scale (BAS) [109];	Family worry of Center for Medical Home Improvement [77];		[77,109]
	Health literacy and knowledge	 Newest Vital Sign UK (NVS-UK) [18]; Diabetes Knowledge Test (DKT) [126]; 	Ad hoc [81,149];	Score of learning test [66,134] *;	[18], [66] *, [81,126], [134] *, [149]
	Compliance/ adherence		- Adaptation from Hausman 2004 [23,146]; - Ad hoc [23];	Tracking completion of indicated quarterly follow-up screens [72];	[23,72,146];
Learning and change behaviour/ attitude	Decisional conflict/ self- determination/ self-advocacy	 Decisional Conflict Scale (dcs) [81]; SURE (Sure of myself, Understand information, Risk-benefit ratio, Encouragement) [81]; Self-Determination Theory (SDT) scales [150]; 	Adaptation from: - Behavioural Regulation in Exercise Questionnaire [151]; - Patient Self-Advocacy Scale [149];		[81,149–151]
	Others	 Recovery Attitudes Questionnaire7 (RAQ-7) [149]; ATT-19 [126]; Physical Activity Questionnaire for adolescent [138]; 	 Unhealthy eating adapted from Stansfeld et al., 2003 [138]; Ad hoc [66,100]*, [138,149]; 	 N. of consultations offered and attended were measured by exercise referral practitioners [49,126]; Job opportunities after or during recovery [135]; 	[49], [66] *, [100] *, [126,135,138,149]
Relationship strengths with professionals			Adaptation from Rajah et al., 2008 [73];		[73]
Issue awareness			Ad hoc [99];		[99]
Cost savings			Cost minimization analysis [16]; Ad hoc health expenses [140];		[16,140]

The asterisk * indicates publications in public domain.

3.3.4. Outcome for Community

The metrics that assess outcomes on the community are the fewest. They have been classified into (i) value for the community and (ii) value for society (Table 9). No metrics have been identified for assessing the third category devised by Cepiku et al. (2020), socioeconomic impact [29].

Value for community has been defined as increasing trust towards service providers and the better understanding of service costs and procedures thanks to direct collaboration with citizens [152]. Although the co-production literature often refers to this concept in explaining outcomes of co-production [2], it is recognised as difficult to define and, especially, to measure. Indeed, only a few studies have used quantitative metrics for assessing value for community regarding social capital. These studies mainly refer to community co-production, for instance in public health interventions, such as Bolton et al. (2016) and Brown et al. (2020), who adopted the Adapted Social Capital Questionnaire [153] and Arizona Social Support Interview Schedule [154], respectively [18,48].

Value for society has been defined as the result of three co-production effects: democratisation of the process, equal distribution of effects on society and increasing public acceptance [29]. The equal distribution of co-production effects on society has been measured with a validated scale, adapted instrument, ad hoc items and indicators. More precisely, it assesses the reduction of health or other public services, for example inequalities in deprived areas or poor countries, for instance, regarding the increasing number of service participants [20,76], the increasing identification of children's needs (National Survey of Children with Special Health Care Needs [155]) and the increasing quality and accessibility of public commodities such as water [64].

In assessing the effects on the community, the public findings bring important insights to the health field, especially in measuring the value for society. While health findings focus on effects related to community health and well-being [77], public ones broaden the perspective by investigating the accessibility and economic affordability of the service and environmental effects.

International Validated Main Other Tools Indicators Outcome References Arizona Social Support Value for Social Adapted Social Capital Interview Schedule (ASSIS) [18] [18,48]community Questionnaire [48] % Increase in participation in the screening program with vulnerable population [20]; National Survey of Increase access to health or [20], [64] *, [76] *, [77], [80] *, [156] Children with Special Ad hoc [64] * Value for society services/ public services [76,156] *: democratization Health Care Needs [77]: Count variable calculating by summing the pro-environmental activities undertaken [80] *

Table 9. Measures and metrics of outcomes for community.

The asterisk * indicates publications in public domain.

4. Discussion

The review confirms that healthcare is a sector where the interest in co-production evaluation is significantly increasing [1–3]. Health is also a highly investigated field among public sector scholars, where the urgency of focusing a co-production research agenda on outcome evaluation is widely recognised [15].

Despite this increasing attention of academics and practitioners, the literature on co-production empirical studies returned a fragmented picture regarding study design (defining "when" co-production outcomes are assessed), approaches, methods and tools (identifying "how" is measured) and specific metrics adopted ("what" is measured).

Considering the research design, less than a quarter of the included studies are based on longitudinal or mixed design (i.e., cross-sectional and longitudinal). It implies that most target the hic and nunc relationships between features of co-production and outcomes, without comparing groups (e.g., co-producers vs. non-co-producers), and ex ante and ex post benchmarks or how outcomes evolve over time (i.e., co-production sustainability). The adoption of a comparative research design would help to enforce the robustness and sustainability of the results of the evaluation effort.

Moreover, the analysis of the research approach reveals the clear predominance of qualitative studies, while the use of quantitative or mixed methods is limited to a quarter of the included papers. Considering the complexity of the topic and its multi-stakeholder and intrinsically multilevel nature, the combination of quantitative and qualitative approaches would make its understanding more robust and valid [157].

Delving deeper into the quantitative approach, findings present a wide range of tools and metrics (i.e., internationally validated scales, other reliable tools, adapted or ad hoc instruments, or single indicators), used for the evaluation of co-production. The study also provides their classification by specific outcome and actor. The paper offers a multidimensional measurement system that researchers and/or managers can use by choosing the most suitable tools, according to the specific measurement needs (e.g., type of co-production, context, actor). For instance, in the case of e-health co-design, as in theextant literature (e.g., [69]), it may be useful to use the SUS scale and measure the acceptability of the co-created e-tool. Metrics aimed at measuring changes in skills and personal behaviour in clinical routines, such as professionals' empathy or attitude towards patients, may help to evaluate the outcomes on co-delivery training (as in [93]). Furthermore, the measure of improvement (or worsening) of health status has often been used in assessing the effects of a co-delivered health service (e.g., [50]); however, the choice of the most appropriate measure should also take into account the specific health context/disease of the patient (e.g., diabetes, motor difficulties, mental disability, etc.).

Currently, some important differences have emerged throughout the actors considered (i.e., regular providers, professionals, lay actors and community) [29].

Specifically, the use of validated tools, borrowed from medicine (e.g., General Health Questionnaire-12 [48]), health psychology (e.g., Hospital Anxiety and Depression Scale (HADS) [119,120,136,137]) and organisational well-being studies (e.g., Social production function scale (SPF-IL) [46,88]), is more widely adopted to evaluate the impacts on lay actors and professionals. Recognised and validated constructs and scales have been used mainly in experimental health research and health service management research, while scarce contributions originated from the public sector. Such results could be fascinating for scholars investigating impacts for users of public service and public officials.

Looking closely at outcome dimensions, unsurprisingly, lay actors are the perspective that captured the main attention of scholars in the health field, with almost half of the retrieved outcomes. The number of sub-dimensions (15) also reflects the maturity of evaluation on this topic. Health status, well-being and quality of life are largely the dimensions that offer a portfolio of validated scales (e.g., health status: Patient Health Questionnaire (PHQ-9) [18]; well-being: Warwick-Edinburgh Mental Well-being Scale (WEMWBS); [48,109,135,136]; quality of life: EuroQol index (EQ 5D index) [119,120]), followed by satisfaction with service (e.g., Psychological Needs Satisfaction in Exercise Scale [49]). Interestingly, some well-known constructs in medicine and health psychology (such as enjoyment [43], self-management [119,120], self-efficacy [50], self-esteem [110], eustress [23], burden [109], trust [73] and issue awareness [99]) have been relatively scarcely analysed.

Overall, the review returns a paucity of attention to a key player in health co-production: the informal caregiver (e.g., [77,110,116]). Given the scant investigation of informal caregiver perspectives, further research is needed to explore the outcome dimensions related to the stakeholders and the research approach that best can evaluate their role. Helping with the daily management of an illness causes caregivers to display a predictable de-

crease in psycho-physical well-being; indeed, the literature describes them as "hidden patients" [158]. As claimed by Ens et al., (2014), the consequences of caregivers' inadequate support negatively influence not only the patient, but also the whole healthcare system (e.g., may cause a greater dependency on institutional healthcare providers and increase medical visits needs, also impacting overall healthcare costs) [159]. Thus, the inadequacy of the caregiver could also act as a deterrent in co-producing.

Last, it is interesting to call attention to the ad hoc scale suggested by Chen et al., (2015) in the public realm, which also aims at explicitly investigating possible negative changes of behaviours in lay actors [100].

The regular provider perspective is the second-most investigated outcome after lay actors. It has been mainly investigated through ad hoc tools or measures, except cost savings (thanks to the well-developed cost-benefits analysis [160]) and more technical service conditions (such as usability and acceptability/feasibility, e.g., [48,69,71,81,82]). In this domain, the public sector literature offers interesting hints with regard to the service improvement dimension (assessing frequency, distance from users [64] rather than the mere number of service users involved [20,76]) and the reorganisation of the government structure regarding the decentralisation of power [76].

Only 17% of the investigated outcomes considered the professionals' perspective. Staff well-being and empathy/attitude are more widely analysed (e.g., [46,88,93,94,101]), thanks to well-known international scales targeting these issues. Public metrics assessing the type of outcome are limited in number and do not add valuable and innovative insights to health findings (e.g., [100]).

Finally, in assessing outcomes on community, the contribution from the health literature is still limited (e.g., [18,77]). The public findings bring important insights to the health field, especially in measuring the value for society. While health findings focus mainly on effects related to community health and well-being [77], public ones broaden the perspective by examining the accessibility [76,156] and economic affordability of the service [64] and environmental effects [80].

5. Conclusions

The results presented and discussed in this article provide a blueprint co-production multidimensional performance measurement system that factors in the values and perspectives of multiple stakeholders. The coding framework has been structured according to the outcome dimensions in the co-production literature [15,29], adapting the analytical components concerning health specificities. In this way, the findings guide the further development of a theoretical framework of value co-production. The paper offers a systematisation of the methods, tools and metrics used to assess outcomes of co-production. Thus far, studies have discussed the effects of co-production, but they have not focused on how these outcomes have been evaluated. Moreover, the framework developed in this paper identifies a new stakeholder dimension: professionals. The role of the workforce is particularly interesting to evaluate separately from the regular provider, as a required specific approach.

This framework can be used in other contexts with cross-sector adaptability. Health emerged as a widely investigated field for the assessment of co-production; this health literature embraces both public and private streams. Health resulted in the top investigated domain in co-production public management; moreover, the service management literature has increasingly focused on the healthcare service setting, considered a rich and fertile context in which to explore a new service delivery model due to its uncommon and complex characteristics [27].

Thus, the findings offer a blueprint multidimensional quantitative performance measurement system that can inform evaluation co-production academic research across research fields, healthcare, public and service management.

The results offer a systematic overview of the literature focusing on co-production evaluation in health, providing performance measurement system tools and measures

available and tested in the literature for each sub-dimension. These indications help to strengthen the understanding and analysis of co-production outcomes. The performance system promotes improvements in empirical data collection on the multiple faceted co-produced activities and spurs consciousness of the adoption of sustainable patient-based initiatives.

This paper offers the opportunity to develop a multi-metric evaluation study. The results showed that the current research is mainly focused on mono-stakeholder impact. This study provides the opportunity to design a configurational approach that goes beyond the focus on a single outcome to illuminate how multiple outcomes interact and interrelate. Outcomes can self-reinforce each other (e.g., the enhancement in professional motivation and patient engagement) or manifest trade-offs (e.g., provider costs or efforts and patient satisfaction).

Finally, the paper offers insights for managers and community leaders engaged in co-production initiatives, supplying a conceptual classification of several impacts affecting co-production and an operational guide to better design and implement an empirical multi-dimensional performance management system. COVID-19 has forced an increase in the co-production of services in the public and health sectors [11]. This trend will be consolidated to boost the development of new co-production practices that will require a robust and replicable measurement and evaluation system.

Noticeably, the paper has also revealed five critical research gaps that could inform the further research agenda on co-production. First, there is a lack of consideration of the impact of the co-production process on informal caregivers and their crucial role in sustaining or interfering with patients' behaviours and motivation to co-produce. Secondly, another understudied area is the impact of co-production on professionals. As workers' wellbeing is an important quality indicator in the delivery of high-quality care, it is important to investigate if and how co-producing with the patient can lead to an increase or decrease in the professionals' wellbeing, including relevant individual and organisational variables. Third, community is not only a neglected outcome dimension, but it also suffers from weaker methodological approaches. Furthermore, the literature still uncovers unfavourable effects of co-production across the multi-stakeholder outcome dimension. Last, the findings still show a lack of empirical mixed-method approaches applied to co-production, which could enhance with robust and valid insights the understanding of how co-produced service affects various stakeholders. A deeper investigation of the type and robustness of qualitative approaches could help to select a more suitable research tool to address the evaluation effort. Moreover, a longitudinal perspective would strengthen the understanding of how co-production impacts unfold over time, addressing the sustainability challenge.

Although the paper aims to be as comprehensive and replicable as possible, it contains some limitations, mostly directly related to the systematic review method. The choice of keywords and the formulation of the query used in search strategy may have resulted in some eligible studies being undetected because they used a different terminology. However, including both co-production and co-creation, which are often used interchangeably [33,34], as well as resorting to the snowball technique, has helped to mitigate this issue. The second limitation is caused by the inclusion of only peer-reviewed journal publications in English. Despite this choice ensuring a scientific and methodological rigour of the findings, it neglected books, conference papers, grey literature and works published in other languages which could have added other interesting insights. Finally, some evaluation approach may be missed because the paper aims at collecting empirical evidence and the theoretical papers and protocols were excluded.

Despite these limitations, the review presents points of originality and was carried out on a large number of studies with a transparent and well-documented process. It helps to advance knowledge of co-production and supports researchers and practitioners in future endeavours.

Supplementary Materials: The following are available online at https://www.mdpi.com/1660-4 601/18/7/3336/s1. Table S1: Main journal and subject area of reviewed articles; Table S2: Most analysed countries in the reviewed articles; Table S3: Study context in public domain.

Author Contributions: The conceptual structure of this paper has been outlined by M.M. and C.G. Based on such a conceptual structure, M.M. led the development of the introduction section, the research aims/methods and the discussions; F.F. performed the literature review on health and E.G. on public literature, writing the first draft of the results session, that has then been reviewed and finalized by M.M. and C.G. Conclusions have been written by M.M. and C.G. All authors have read and agreed to the published version of the manuscript.

Funding: No funding was obtained for this study.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available on request.

Acknowledgments: The authors acknowledge the support from the University of Milan through APC initiative.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Dunston, R.; Lee, A.; Boud, D.; Brodie, P.; Chiarella, M. Co-production and health system reform—From re-imagining to re-making. *Aust. J. Public Adm.* **2009**, *68*, 39–52. [CrossRef]
- 2. Voorberg, W.H.; Bekkers, V.J.J.M.; Tummers, L.G. A Systematic Review of Co-Creation and Co-Production: Embarking on the social innovation journey. *Public Manag. Rev.* **2015**, *17*, 1333–1357. [CrossRef]
- 3. McMullin, C.; Needham, C. Co-production in Healthcare. In *Co-Production and Co-Creation: Engaging Citizens in Public Services*; Brandsen, T., Steen, T.P.S., Verschuere, B., Eds.; Routledge: New York, NY, USA, 2018; pp. 151–159. ISBN 9781138700116.
- 4. Department of Health and Social Care. *Our Health, Our Care, Our Say: A New Direction for COMMUNITY Services*; Department of Health and Social Care: Norwich, UK, 2006.
- 5. Trummer, U.F.; Mueller, U.O.; Nowak, P.; Stidl, T.; Pelikan, J.M. Does physician-patient communication that aims at empowering patients improve clinical outcome?. A case study. *Patient Educ. Couns.* **2006**, *61*, 299–306. [CrossRef]
- 6. Ding, B.; Liu, W.; Tsai, S.B.; Gu, D.; Bian, F.; Shao, X. Effect of patient participation on nurse and patient outcomes in inpatient healthcare. *Int. J. Environ. Res. Public Health* **2019**, *16*, 1344. [CrossRef]
- 7. World Heath Organisation. Facing the Future: Opportunities and Challenges for 21st-Century Public Health in Implementing the Sustainable Development Goals and the Health 2020 Policy Framework; World Heath Organisation: Geneva, Switzerland, 2018.
- 8. Hollander, J.E.; Brendan, G.; Carr, M.D. Virtually Perfect? Telemedicine for Covid-19. *N. Engl. J. Med.* **2020**, *382*, 1679–1681. [CrossRef] [PubMed]
- 9. Colombo, C.; Burgel, G.; Gartner, S.; van Koningsbruggen-Rietschel, S.; Naehrlich, L.; Sermet-Gaudelus, I.; Southern, K.W. Impact of COVID-19 on people with cystic fibrosis. *Lancet Respir.* **2020**, *8*, 35–36. [CrossRef]
- 10. Cepiku, D.; Giordano, F.; Bovaird, T.; Loeffler, E. New development: Managing the Covid-19 pandemic—From a hospital-centred model of care to a community co-production approach. *Public Money Manag.* **2021**, *41*, 77–80. [CrossRef]
- 11. Steen, T.; Brandsen, T. Coproduction during and after the COVID-19 Pandemic: Will It Last? *Public Adm. Rev.* **2020**, *80*, 851–855. [CrossRef] [PubMed]
- 12. Tuurnas, S.P.; Stenvall, J.; Rannisto, P.H.; Harisalo, R.; Hakari, K. Coordinating co-production in complex network settings. *Eur. J. Soc. Work* **2015**, *18*, 370–382. [CrossRef]
- 13. Bovaird, T. Beyond engagement and participation: User and community coproduction of public services. *Public Adm. Rev.* **2007**, 67, 846–860. [CrossRef]
- 14. Vainieri, M.; Noto, G.; Ferre, F.; Rosella, L.C. A performance management system in healthcare for all seasons? *Int. J. Environ. Res. Public Health* **2020**, *17*, 5590. [CrossRef] [PubMed]
- 15. Loeffler, E. Co-Production of Public Services and Outcomes; Palgrave Macmillan: London, UK, 2020.
- 16. Spanò, R.; Di Paola, N.; Bova, M.; Barbarino, A. Value co-creation in healthcare: Evidence from innovative therapeutic alternatives for hereditary angioedema. *BMC Health Serv. Res.* **2018**, *18*, 1–13. [CrossRef] [PubMed]
- 17. De Rosis, S.; Pennucci, F.; Noto, G.; Nuti, S. Healthy living and co-production: Evaluation of processes and outcomes of a health promotion initiative co-produced with adolescents. *Int. J. Environ. Res. Public Health* **2020**, *17*, 8007. [CrossRef]
- 18. Brown, J.; Luderowski, A.; Namusisi-Riley, J.; Moore-Shelley, I.; Bolton, M.; Bolton, D. Can a community-led intervention offering social support and health education improve maternal health? A repeated measures evaluation of the pact project run in a socially deprived london borough. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2759. [CrossRef] [PubMed]

- 19. Gallan, A.S.; Jarvis, C.B.; Brown, S.W.; Bitner, M.J. Customer positivity and participation in services: An empirical test in a health care context. *J. Acad. Mark. Sci.* **2013**, *41*, 338–356. [CrossRef]
- 20. Eriksson, E.M. Representative co-production: Broadening the scope of the public service logic. *Public Manag. Rev.* **2019**, 21, 291–314. [CrossRef]
- 21. Sweeney, J.C.; Danaher, T.S.; McColl-Kennedy, J.R. Customer Effort in Value Cocreation Activities: Improving Quality of Life and Behavioral Intentions of Health Care Customers. *J. Serv. Res.* **2015**, *18*, 318–335. [CrossRef]
- 22. Pham, T.A.N.; Sweeney, J.C.; Soutar, G.N. Customer value cocreation activities: An exploration of psychological drivers and quality of life outcomes. *J. Serv. Theory Pract.* **2019**, 29, 282–308. [CrossRef]
- 23. Mende, M.; Scott, M.L.; Bitner, M.J.; Ostrom, A.L. Activating consumers for better service coproduction outcomes through eustress: The interplay of firm-assigned workload, service literacy, and organizational support. *J. Public Policy Mark.* **2017**, *36*, 137–155. [CrossRef]
- 24. Jaspers, S.; Steen, T. The sustainability of outcomes in temporary co-production. *Int. J. Public Sect. Manag.* **2019**, *33*, 62–77. [CrossRef]
- 25. Fusco, F.; Marsilio, M.; Guglielmetti, C. Co-production in health policy and management: A comprehensive bibliometric review. *BMC Health Serv. Res.* **2020**, *20*, 1–16. [CrossRef]
- 26. Dudau, A.; Glennon, R.; Verschuere, B. Following the yellow brick road? (Dis)enchantment with co-design, co-production and value co-creation in public services. *Public Manag. Rev.* **2019**, 21, 1577–1594. [CrossRef]
- 27. McColl-Kennedy, J.R.; Snyder, H.; Elg, M.; Witell, L.; Helkkula, A.; Hogan, S.J.; Anderson, L. The changing role of the health care customer: Review, synthesis and research agenda. *J. Serv. Manag.* **2017**, *28*, 2–33. [CrossRef]
- 28. Williams, B.N.; Kang, S.C.; Johnson, J. (Co)-Contamination as the Dark Side of Co-Production: Public value failures in co-production processes. *Public Manag. Rev.* **2016**, *18*, 692–717. [CrossRef]
- 29. Cepiku, D.; Marsilio, M.; Sicilia, M.; Vainieri, M. *The Co-Production of Public Services Management and Evaluation*; Palgrave Macmillan: Cham, Switzerland, 2020; ISBN 9783030607098.
- 30. Moher, D.; Shamseer, L.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A.; Group, P. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst. Rev.* **2015**, *4*, 1–9. [CrossRef]
- 31. Moher, D.; Liberati, A.; Tetzlaff, J.; Altman, D.G.; Altman, D.; Antes, G.; Atkins, D.; Barbour, V.; Barrowman, N.; Berlin, J.A.; et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med.* **2009**, *6*. [CrossRef]
- 32. Nabatchi, T.; Sancino, A.; Sicilia, M. Varieties of Participation in Public Services: The Who, When, and What of Coproduction. *Public Adm. Rev.* **2017**, *77*, 766–776. [CrossRef]
- 33. Vargo, S.L.; Lusch, R.F. Service-dominant logic: Continuing the evolution. J. Acad. Mark. Sci. 2008, 36, 1–10. [CrossRef]
- 34. Vargo, S.L.; Lusch, R.F. Evolving to a new dominant logic for marketing. *Serv. Log. Mark. Dialog, Debate, Dir.* **2004**, *68*, 3–28. [CrossRef]
- 35. Gebauer, H.; Johnson, M.; Enquist, B. Value co-creation as a determinant of success in public transport services: A study of the Swiss Federal Railway operator (SBB). *Manag. Serv. Qual.* **2010**, *20*, 511–530. [CrossRef]
- 36. Loeffler, E.; Bovaird, T. From participation to co-production: Widening and deepening the contributions of citizens to public services and outcomes. In *The Palgrave Handbook of Public Administration and Management in Europe*; Ongaro, E., Van Thiel, S., Eds.; Palgrave Macmillan UK: London, UK, 2017; pp. 403–423. ISBN 9781137552693.
- 37. Tranfield, D.; Denyer, D.; Smart, P. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *Br. J. Manag.* **2003**, *14*, 207–222. [CrossRef]
- 38. Lega, F.; Marsilio, M.; Villa, S. An evaluation framework for measuring supply chain performance in the public healthcare sector: Evidence from the Italian NHS. *Prod. Plan. Control* **2013**, 24, 931–947. [CrossRef]
- 39. Batalden, M.; Batalden, P.; Margolis, P.; Seid, M.; Armstrong, G.; Opipari-Arrigan, L.; Hartung, H. Coproduction of healthcare service. *BMJ Qual. Saf.* **2016**, *25*, 509–517. [CrossRef]
- 40. Wallace, J.E.; Lemaire, J.B.; Ghali, W.A. Physician wellness: A missing quality indicator. Lancet 2009, 374, 1714–1721. [CrossRef]
- 41. Parkinson, M.D. The Healthy Health Care Workplace: A Competitive Advantage. *Curr. Cardiol. Rep.* **2018**, 20. [CrossRef] [PubMed]
- 42. Sehgal, S.; Gupta, G. Converging resources and co-producing for innovation: Evidence from healthcare services. *Eur. J. Innov. Manag.* **2019**, 23, 429–453. [CrossRef]
- 43. Hau, L.N. The role of customer operant resources in health care value creation. Serv. Bus. 2018. [CrossRef]
- 44. McColl-Kennedy, J.R.; Hogan, S.J.; Witell, L.; Snyder, H. Cocreative customer practices: Effects of health care customer value cocreation practices on well-being. *J. Bus. Res.* **2017**, *70*, 55–66. [CrossRef]
- 45. Kim, J. Customers' value co-creation with healthcare service network partners: The moderating effect of consumer vulnerability. *J. Serv. Theory Pract.* **2019**, 29, 309–328. [CrossRef]
- 46. Den Boer, J.; Nieboer, A.P.; Cramm, J.M. A cross-sectional study investigating patient-centred care, co-creation of care, well-being and job satisfaction among nurses. *J. Nurs. Manag.* **2017**, *25*, 577–584. [CrossRef]
- 47. Lamph, G.; Sampson, M.; Smith, D.; Williamson, G.; Guyers, M. Can an interactive e-learning training package improve the understanding of personality disorder within mental health professionals? *J. Ment. Health Training Educ. Pract.* **2018**, *13*, 124–134. [CrossRef]

- 48. Bolton, M.; Moore, I.; Ferreira, A.; Day, C.; Bolton, D. Community organizing and community health: Piloting an innovative approach to community engagement applied to an early intervention project in south London. *J. Public Health* **2016**, *38*, 115–121. [CrossRef] [PubMed]
- 49. Buckley, B.J.; Thijssen, D.H.; Murphy, R.C.; Graves, L.E.; Cochrane, M.; Gillison, F.; Crone, D.; Wilson, P.M.; Whyte, G.; Watson, P.M. Pragmatic evaluation of a coproduced physical activity referral scheme: A UK quasi-experimental study. *BMJ Open* **2020**, 10, e034580. [CrossRef]
- 50. Fors, A.; Ekman, I.; Taft, C.; Björkelund, C.; Frid, K.; Larsson, M.E.H.; Thorn, J.; Ulin, K.; Wolf, A.; Swedberg, K. Person-centred care after acute coronary syndrome, from hospital to primary care—A randomised controlled trial. *Int. J. Cardiol.* **2015**, 187, 693–699. [CrossRef]
- 51. Kuipers, S.J.; Cramm, J.M.; Nieboer, A.P. The importance of patient-centered care and co-creation of care for satisfaction with care and physical and social well-being of patients with multi-morbidity in the primary care setting. *BMC Health Serv. Res.* **2019**, 19, 1–9. [CrossRef] [PubMed]
- 52. Grundy, A.C.; Walker, L.; Meade, O.; Fraser, C.; Cree, L.; Bee, P.; Lovell, K.; Callaghan, P. Evaluation of a co-delivered training package for community mental health professionals on service user- and carer-involved care planning. *J. Psychiatr. Ment. Health Nurs.* 2017, 24, 358–366. [CrossRef]
- 53. Dalgarno, M.; Oates, J. The meaning of co-production for clinicians: An exploratory case study of Practitioner Trainers in one Recovery College. *J. Psychiatr. Ment. Health Nurs.* **2018**, 25, 349–357. [CrossRef] [PubMed]
- 54. Farr, M.; Banks, J.; Edwards, H.B.; Northstone, K.; Bernard, E.; Salisbury, C.; Horwood, J. Implementing online consultations in primary care: A mixed-method evaluation extending normalisation process theory through service co-production. *BMJ Open* **2018**, *8*, 1–11. [CrossRef]
- 55. Pocobello, R.; el Sehity, T.; Negrogno, L.; Minervini, C.; Guida, M.; Venerito, C. Comparison of a co-produced mental health service to traditional services: A co-produced mixed-methods cross-sectional study. *Int. J. Ment. Health Nurs.* **2020**, 29, 460–475. [CrossRef]
- 56. Neech, S.G.B.; Scott, H.; Priest, H.M.; Bradley, E.J.; Tweed, A.E. Experiences of user involvement in mental health settings: User motivations and benefits. *J. Psychiatr. Ment. Health Nurs.* **2018**, 25, 327–337. [CrossRef]
- 57. Pinho, N.; Beirão, G.; Patrício, L.; Fisk, R.P. Understanding value co-creation in complex services with many actors. *J. Serv. Manag.* **2014**, 25, 470–493. [CrossRef]
- 58. Beirão, G.; Patrício, L.; Fisk, R.P. Value cocreation in service ecosystems: Investigating health care at the micro, meso, and macro levels. *J. Serv. Manag.* **2017**, *28*, 227–249. [CrossRef]
- 59. Sharma, S.; Conduit, J.; Rao Hill, S. Hedonic and eudaimonic well-being outcomes from co-creation roles: A study of vulnerable customers. *J. Serv. Mark.* **2017**, *31*, 397–411. [CrossRef]
- 60. Flemig, S.; Osborne, S. The Dynamics of Co-Production in the Context of Social Care Personalisation: Testing Theory and Practice in a Scottish Context. *J. Soc. Policy* **2019**, *48*, 1–27. [CrossRef]
- 61. Mayer, C.; McKenzie, K. It shows that there's no limits': The psychological impact of co-production for experts by experience working in youth mental health. *Health Soc. Care Community* **2017**, 25, 1181–1189. [CrossRef]
- 62. Tchanturia, K.; Smith, K.; Glennon, D.; Burhouse, A. Towards an Improved Understanding of the Anorexia Nervosa and Autism Spectrum Comorbidity: PEACE Pathway Implementation. *Front. Psychiatry* **2020**, *11*, 1–7. [CrossRef] [PubMed]
- 63. Bourne, P.; Meddings, S.; Whittington, A. An evaluation of service use outcomes in a Recovery College. *J. Ment. Health* **2018**, 27, 359–366. [CrossRef] [PubMed]
- 64. Adams, E.A.; Boateng, G.O. Are urban informal communities capable of co-production? The influence of community–public partnerships on water access in Lilongwe, Malawi. *Environ. Urban.* **2018**, *30*, 461–480. [CrossRef]
- 65. Park, E.S. Coproduction in the treatment of substance use disorder and its relationship to clinics' service output patterns. *Soc. Serv. Rev.* **2020**, 94, 607–645. [CrossRef]
- 66. Sanina, A.; Kutergina, E.; Balashov, A. The Co-Creative approach to digital simulation games in social science education. *Comput. Educ.* **2020**, *149*, 103813. [CrossRef]
- 67. Jakobsen, M.; Andersen, S.C. Coproduction and Equity in Public Service Delivery. Public Adm. Rev. 2013, 73, 704–713. [CrossRef]
- 68. Brooke, J. SUS: A "Quick and Dirty" Usability Scale. In *Usability Evaluation In Industry*; CRC Press: Boca Raton, FL, USA, 1995; pp. 207–212.
- 69. Amann, J.; Fiordelli, M.; Brach, M.; Bertschy, S.; Scheel-Sailer, A.; Rubinelli, S. Co-designing a self-management app prototype to support people with spinal cord injury in the prevention of pressure injuries: Mixed methods study. *JMIR mHealth uHealth* **2020**, 8, 1–17. [CrossRef] [PubMed]
- 70. Shoemakera, S.J.; Wolfb, M.S.; Brach, C. Development of the Patient Education Materials Assessment Tool (PEMAT): A new measure of understandability and actionability for print and audiovisual patient information. *Patient Educ. Couns.* **2014**, 96, 395–403. [CrossRef] [PubMed]
- 71. Badiu, C.; Bonomi, M.; Borshchevsky, I.; Cools, M.; Craen, M.; Ghervan, C.; Hauschild, M.; Hershkovitz, E.; Hrabovszky, E.; Juul, A.; et al. Developing and evaluating rare disease educational materials co-created by expert clinicians and patients: The paradigm of congenital hypogonadotropic hypogonadism. *Orphanet J. Rare Dis.* 2017, 12, 1–9. [CrossRef]

- 72. Liu, F.F.; Lew, A.; Andes, E.; McNamara, S.; Cassidy, J.; Whitmore, S.; Plunkett, R.; Ong, T. Implementation strategies for depression and anxiety screening in a pediatric cystic fibrosis center: A quality improvement project. *Pediatr. Pulmonol.* **2020**, 55, 3328–3336. [CrossRef]
- 73. Banyte, J.; Tarute, A.; Taujanskyte, I. Customer engagement into value creation: Determining factors and relations with loyalty. *Eng. Econ.* **2014**, *25*, 568–577. [CrossRef]
- 74. Fledderus, J. Does User Co-Production of Public Service Delivery Increase Satisfaction and Trust? Evidence From a Vignette Experiment. *Int. J. Public Adm.* **2015**, *38*, 642–653. [CrossRef]
- 75. Zainuddin, N.; Russell-Bennett, R.; Previte, J. The value of health and wellbeing: An empirical model of value creation in social marketing. *Eur. J. Mark.* **2013**, *47*, 1504–1524. [CrossRef]
- 76. Dhirathiti, N.S. Co-production and the provision of lifelong learning policy for elderly people in Thailand. *Public Manag. Rev.* **2019**, *21*, 1011–1028. [CrossRef]
- 77. McAllister, J.W.; Keehn, R.M.N.; Rodgers, R.; Lock, T.M. Care Coordination Using a Shared Plan of Care Approach: From Model to Practice. *J. Pediatr. Nurs.* **2018**, 43, 88–96. [CrossRef] [PubMed]
- 78. Bentzen, T.Ø. Continuous co-creation: How ongoing involvement impacts outcomes of co-creation. *Public Manag. Rev.* **2020**, 1–21. [CrossRef]
- 79. Landi, S.; Russo, S. Co-production 'thinking' and performance implications in the case of separate waste collection. *Public Manag. Rev.* **2020**, 1–25. [CrossRef]
- 80. Alonso, J.M.; Andrews, R.; Clifton, J.; Diaz-Fuentes, D. Factors influencing citizens' co-production of environmental outcomes: A multi-level analysis. *Public Manag. Rev.* **2019**, *21*, 1620–1645. [CrossRef]
- 81. Bashir, N.Y.; Moore, J.E.; Buckland, D.; Rodrigues, M.; Tonelli, M.; Thombs, B.D.; Bell, N.R.; Isaranuwatchai, W.; Peng, T.; Shilman, D.M.; et al. Are patient education materials about cancer screening more effective when co-created with patients? A qualitative interview study and randomized controlled trial. *Curr. Oncol.* 2019, 26, 124–136. [CrossRef]
- 82. Timmerman, J.G.; Tönis, T.M.; Dekker-Van Weering, M.G.H.; Stuiver, M.M.; Wouters, M.W.J.M.; Van Harten, W.H.; Hermens, H.J.; Vollenbroek-Hutten, M.M.R. Co-creation of an ICT-supported cancer rehabilitation application for resected lung cancer survivors: Design and evaluation. *BMC Health Serv. Res.* **2016**, *16*, 1–11. [CrossRef]
- 83. Arevian, A.C.; O'Hora, J.; Jones, F.; Mango, J.; Jones, L.; Williams, P.G.; Booker-Vaughns, J.; Jones, A.; Pulido, E.; Banner-Jackson, D.; et al. Participatory technology development to enhance community resilience. *Ethn. Dis.* **2018**, *28*, 493–502. [CrossRef]
- 84. Emaldi, M.; Aguilera, U.; López-de-Ipiña, D.; Pérez-Velasco, J. Towards citizen co-created public service apps. *Sensors* **2017**, 17, 1265. [CrossRef] [PubMed]
- 85. Brooks, H.; Irmansyah, I.; Susanti, H.; Utomo, B.; Prawira, B.; Iskandar, L.; Colucci, E.; Keliat, B.A.; James, K.; Bee, P.; et al. Evaluating the acceptability of a co-produced and co-delivered mental health public engagement festival: Mental Health Matters, Jakarta, Indonesia. *Res. Involv. Engagem.* 2019, 5, 1–10. [CrossRef]
- 86. Stallard, P.; Porter, J.; Grist, R. A smartphone app (Blueice) for young people who self-harm: Open phase 1 pre-post trial. *JMIR mHealth uHealth* **2018**, 6. [CrossRef]
- 87. Lee, D.H. Effects of key value co-creation elements in the healthcare system: Focusing on technology applications. *Serv. Bus.* **2019**, 13, 389–417. [CrossRef]
- 88. Van der Meer, L.; Nieboer, A.P.; Finkenflügel, H.; Cramm, J.M. The importance of person-centred care and co-creation of care for the well-being and job satisfaction of professionals working with people with intellectual disabilities. *Scand. J. Caring Sci.* **2018**, 32, 76–81. [CrossRef] [PubMed]
- 89. Breidbach, C.F.; Antons, D.; Salge, T.O. Seamless Service? On the Role and Impact of Service Orchestrators in Human-Centered Service Systems. *J. Serv. Res.* **2016**, *19*, 458–476. [CrossRef]
- 90. Traynor, M.; Wade, B. The development of a measure of job satisfaction for use in monitoring the morale of community nurses in four trusts. *J. Adv. Nurs.* **1993**, *18*, 127–136. [CrossRef]
- 91. Nieboer, A.; Lindenberg, S.; Boomsma, A.; Van Bruggen, A.C. Dimensions of well-being and their measurement: The Spf-II scale. *Soc. Indic. Res.* **2005**, *73*, 313–353. [CrossRef]
- 92. Maslach, C.; Jackson, S.E. The measurement of experienced burnout. J. Organ. Behav. 1981, 2, 99–113. [CrossRef]
- 93. Finamore, C.; Rocca, F.; Parker, J.; Blazdell, J.; Dale, O. The impact of a co-produced personality disorder training on staff burnout, knowledge and attitudes. *Ment. Health Rev. J.* **2020**, 25, 269–280. [CrossRef]
- 94. Farnese, M.L.; Girardi, G.; Fida, R.; Bivona, U.; Bartolo, M.; De Tanti, A.; Intiso, D.; Scarponi, F.; Antonucci, G. Caregivers' engagement during in-hospital care of sABI's patients: Evaluation of informal co-production from the health providers' perspective. *Health Soc. Care Community* **2020**, *28*, 2086–2094. [CrossRef] [PubMed]
- 95. Horne, S.; Hastings, R.P. Positive perceptions held by support staff in community mental retardation services. *Am. J. Ment. Retard.* **2004**, *109*, 53–61. [PubMed]
- 96. Schaufeli, W.B.; Salanova, M.; Gonzalez-Roma, V.; Bakker, A.B. The Measurement Of Engagement And Burnout: A Two Sample Confirmatory Factor Analytic Approach. *J. Happiness Stud.* **2001**, *3*, 71–92. [CrossRef]
- 97. Hastings, R.P.; Gillespie, D.; Flynn, S.; McNamara, R.; Taylor, Z.; Knight, R.; Randell, E.; Richards, L.; Moody, G.; Mitchell, A.; et al. Who's challenging who training for staff empathy towards adults with challenging behaviour: Cluster randomised controlled trial. *J. Intellect. Disabil. Res.* **2018**, *62*, 798–813. [CrossRef] [PubMed]

- 98. Dickens, G.L.; Lamont, E.; Mullen, J.; MacArthur, N.; Stirling, F.J. Mixed-methods evaluation of an educational intervention to change mental health nurses' attitudes to people diagnosed with borderline personality disorder. *J. Clin. Nurs.* **2019**, *28*, 2613–2623. [CrossRef]
- 99. Jo, S.; Nabatchi, T. Coproducing healthcare: Individual-level impacts of engaging citizens to develop recommendations for reducing diagnostic error. *Public Manag. Rev.* **2019**, *21*, 354–375. [CrossRef]
- 100. Chen, C.M.; Lee, P.C.; Chou, C.H. The Impact of Coproductive Taxpayers' Supervisory Behaviors on the Job Involvement of Tax Collectors. *Rev. Public Pers. Adm.* **2015**, 35, 278–304. [CrossRef]
- 101. Davies, J.; Sampson, M.; Beesley, F.; Smith, D.; Baldwin, V. An evaluation of Knowledge and Understanding Framework personality disorder awareness training: Can a co-production model be effective in a local NHS mental health Trust? *Personal. Ment. Health* 2014, 7, 161–168. [CrossRef]
- 102. Horgan, A.; Manning, F.; Bocking, J.; Happell, B.; Lahti, M.; Doody, R.; Griffin, M.; Bradley, S.K.; Russell, S.; Bjornsson, E.; et al. 'To be treated as a human': Using co-production to explore experts by experience involvement in mental health nursing education The COMMUNE project. *Int. J. Ment. Health Nurs.* **2018**, 27, 1282–1291. [CrossRef] [PubMed]
- 103. Manning, J.C.; Carter, T.; Latif, A.; Horsley, A.; Cooper, J.; Armstrong, M.; Crew, J.; Wood, D.; Callaghan, P.; Wharrad, H. "Our Care through Our Eyes". Impact of a co-produced digital educational programme on nurses' knowledge, confidence and attitudes in providing care for children and young people who have self-harmed: A mixed-methods study in the UK. *BMJ Open* **2017**, 7. [CrossRef] [PubMed]
- 104. Ryu, B.; Kim, N.; Heo, E.; Yoo, S.; Lee, K.; Hwang, H.; Kim, J.W.; Kim, Y.; Lee, J.; Jung, S.Y. Impact of an electronic health record-integrated personal health record on patient participation in health care: Development and randomized controlled trial of myhealthkeeper. *J. Med. Internet Res.* 2017, 19, 1–12. [CrossRef]
- 105. Kroenke, K.; Spitzer, R.L. The PHQ-9: A new depression diagnostic and severity measure. *Psychiatr. Ann.* **2002**, *32*, 509–5015. [CrossRef]
- 106. Wing, J.K.; Beevor, A.S.; Curtis, R.H.; Park, S.G.B.; Hadden, J.; Burns, A. Health of the Nation Outcome Scales (HoNOS): Research and development. *Br. J. Psychiatry* **1998**, *172*, 11–18. [CrossRef]
- 107. Adinolfi, P.; Starace, F.; Palumbo, R. Health Outcomes and Patient Empowerment: The Case of Health Budgets in Italy Article. *J. Health Manag.* **2016**, *18*, 117–133. [CrossRef]
- 108. Tennant, R.; Hiller, L.; Fishwick, R.; Platt, S.; Joseph, S.; Weich, S.; Parkinson, J.; Secker, J.; Stewart-Brown, S. The Warwick-Dinburgh mental well-being scale (WEMWBS): Development and UK validation. *Health Qual. Life Outcomes* **2007**, *5*, 1–13. [CrossRef]
- 109. Chiocchi, J.; Lamph, G.; Slevin, P.; Fisher-Smith, D.; Sampson, M. Can a carer (peer) led psychoeducation programme improve mental health carers well-being, reduce burden and enrich empowerment: A service evaluation study. *J. Ment. Health Training Educ. Pract.* 2019, 14, 131–140. [CrossRef]
- 110. Wood, T.E.; Englander-Golden, P.; Golden, D.E.; Pillai, V.K. Improving addictions treatment outcomes by empowering self and others. *Int. J. Ment. Health Nurs.* **2010**, *19*, 363–368. [CrossRef] [PubMed]
- 111. Englander-Golden, P.E.; Golden, D.E. *Say it Straight: Training in Straightforward Communication*; Golden Communication Books, Say It Straight Foundation: Carlsbad, CA, USA, 2003.
- 112. Davis, J.R.; Rawana, E.P.; Capponi, D.R. Acceptability of behavioral staff management techniques. In *Behavioral Interventions* 4.1; John Wiley & Sons Ltd.: Hoboken, NJ, USA, 1989; pp. 22–44.
- 113. Leo, C.; Zainuddin, N. Exploring value destruction in social marketing services. J. Soc. Mark. 2017, 7, 405–422. [CrossRef]
- 114. Hau, L.N.; Thuy, P.N. Impact of service personal values on service value and customer loyalty: A cross-service industry stud. *Serv. Bus.* **2012**, *6*, 137–155. [CrossRef]
- 115. Cramm, J.M.; Strating, M.M.H.; Nieboer, A.P. Validation of the Caregivers' Satisfaction with Stroke Care Questionnaire: C-SASC hospital scale. *J. Neurol.* **2011**, 258, 1008–1012. [CrossRef] [PubMed]
- 116. Hakobyan, L.; Nieboer, A.P.; Finkenflügel, H.; Cramm, J.M. The Significance of Person-Centered Care for Satisfaction With Care and Well-Being Among Informal Caregivers of Persons With Severe Intellectual Disability. *J. Policy Pract. Intellect. Disabil.* **2020**, 17, 31–42. [CrossRef]
- 117. Center of Excellence on Quality of Care Measures for Children with Complex Needs Family Experiences with Care Coordination Measure set (FECC). Available online: https://www.ahrq.gov/sites/default/files/wysiwyg/pqmp/measures/chronic/chipra-166-fullreport.pdf (accessed on 27 September 2020).
- 118. Hibbard, J.H.; Stockard, J.; Mahoney, E.R.; Tusler, M. Development of the patient activation measure (PAM): Conceptualizing and measuring activation in patients and consumers. *Health Serv. Res.* **2004**, *39*, 1005–1026. [CrossRef]
- 119. Turner, A.; Realpe, A.X.; Wallace, L.M.; Kosmala-Anderson, J. A co-produced self-management programme improves psychosocial outcomes for people living with depression. *Ment. Health Rev. J.* **2015**, 20, 242–255. [CrossRef]
- 120. Turner, A.; Anderson, J.K.; Wallace, L.M.; Bourne, C. An evaluation of a self-management program for patients with long-term conditions. *Patient Educ. Couns.* **2015**, *98*, 213–219. [CrossRef]
- 121. Petukiene, E.; Damkuvienė, M. Participation of Clients in Public Services: The Aspect of Educating to Participate. *Soc. Sci.* **2011**, 73, 301–309. [CrossRef]
- 122. Thomsen, M.K.; Jakobsen, M. Influencing Citizen Coproduction by Sending Encouragement and Advice: A Field Experiment. *Int. Public Manag. J.* **2015**, *18*, 286–303. [CrossRef]

- 123. Hyett, N.; Bagley, K.; Iacono, T.; McKinstry, C.; Spong, J.; Landry, O. Evaluation of a Codesign Method Used to Support the Inclusion of Children With Disability in Mainstream Schools. *Int. J. Qual. Methods* **2020**, *19*, 1–12. [CrossRef]
- 124. Osborne, R.H.; Elsworth, G.R.; Whitfield, K. The Health Education Impact Questionnaire (heiQ): An outcomes and evaluation measure for patient education and self-management interventions for people with chronic conditions. *Patient Educ. Couns.* 2007, 66, 192–201. [CrossRef] [PubMed]
- 125. Toobert, D.J.; Hampson, S.E.; Glasgow, R.E. The Summary of Diabetes Self-Care. Diabetes Care J. 2000, 23, 943–950. [CrossRef]
- 126. New, N. Teaching so they hear: Using a co-created diabetes self-management education approach. *J. Am. Acad. Nurse Pract.* **2010**, 22, 316–325. [CrossRef]
- 127. Fors, A.; Gyllensten, H.; Swedberg, K.; Ekman, I. Effectiveness of person-centred care after acute coronary syndrome in relation to educational level: Subgroup analysis of a two-armed randomised controlled trial. *Int. J. Cardiol.* 2016, 221, 957–962. [CrossRef] [PubMed]
- 128. Fors, A.; Blanck, E.; Ali, L.; Ekberg-Jansson, A.; Fu, M.; Kjellberg, I.L.; Mäkitalo, Å.; Swedberg, K.; Taft, C.; Ekman, I. Effects of a person-centred telephone-support in patients with chronic obstructive pulmonary disease and/or chronic heart failure—A randomized controlled trial. *PLoS ONE* **2018**, *13*, 1–12. [CrossRef]
- 129. Johnston, M.; Wright, S.; Weinman, J. Measures in Health Psychology: A User's Portfolio. Causal and Control Beliefs; NFER-NELSON: Windsor, UK, 1995.
- 130. Rosenberg, M. Society and the Adolescent Self-Image; Princeton University Press: Priston, NJ, USA, 1979.
- 131. Reinhard, S.; Gubman, G.; Horwitz, A.; Minsky, S. Burden assessment scale for families of the seriously mentally ill. *Eval. Program Plann.* **1994**, *17*, 261–269. [CrossRef]
- 132. Weiss, B.D.; Mays, M.Z.; Martz, W.; Castro, K.M.; DeWalt, D.A.; Pignone, M.P.; Mockbee, J.; Hale, F.A. Quick assessment of literacy in primary care: The newest vital sign. *Ann. Fam. Med.* **2005**, *3*, 514–522. [CrossRef]
- 133. Fitzgerald, J.T.; Funnell, M.M.; Hess, G.E.; Barr, P.A.; Anderson, R.M.; Hiss, R.G.; Davis, W.K. The reliability and validity of a brief diabetes knowledge test. *Diabetes Care J.* 1998, 21, 706–710. [CrossRef]
- 134. Jakobsen, M. Can government initiatives increase citiz. J. Public Adm. Res. Theory 2013, 23, 27–54. [CrossRef]
- 135. Kay, K.; Edgley, G. Evaluation of a new recovery college: Delivering health outcomes and cost efficiencies via an educational approach. *Ment. Health Soc. Incl.* **2019**, 23, 36–46. [CrossRef]
- 136. Dodd-Reynolds, C.J.; Nevens, L.; Oliver, E.J.; Finch, T.; Lake, A.A.; Hanson, C.L. Prototyping for public health in a local context: A streamlined evaluation of a community-based weight management programme (Momenta), Northumberland, UK. *BMJ Open* **2019**, *9*, 1–11. [CrossRef]
- 137. Hauffman, A.; Alfonsson, S.; Bill-Axelson, A.; Bergkvist, L.; Forslund, M.; Mattsson, S.; von Essen, L.; Nygren, P.; Igelström, H.; Johansson, B. Cocreated internet-based stepped care for individuals with cancer and concurrent symptoms of anxiety and depression: Results from the U-CARE AdultCan randomized controlled trial. *Psychooncology.* 2020, 29, 2012–2018. [CrossRef] [PubMed]
- 138. Frostick, C.; Watts, P.; Netuveli, G.; Renton, A.; Moore, D. Well London: Results of a Community Engagement Approach to Improving Health Among Adolescents from Areas of Deprivation in London. *J. Community Pract.* 2017, 25, 235–252. [CrossRef]
- 139. Weinberg, D.B.; Lusenhop, R.W.; Gittell, J.H.; Kautz, C.M. Providers and Informal Caregivers. *Health Care Manag. Rev.* **2007**, 32, 140–149. [CrossRef]
- 140. Damali, U.; Miller, J.L.; Fredendall, L.D.; Moore, D.W.; Dye, C.J. Co-creating value using customer training and education in a healthcare service design. *J. Oper. Manag.* **2016**, *47*–*48*, 80–97. [CrossRef]
- 141. Heim, N.; Rolden, H.; van Fenema, E.M.; Weverling-Rijnsburger, A.W.E.; Tuijl, J.P.; Jue, P.; Oleksik, A.M.; de Craen, A.J.M.; Mooijaart, S.P.; Blauw, G.J.; et al. The development, implementation and evaluation of a transitional care programme to improve outcomes of frail older patients after hospitalisation. *Age Ageing* **2016**, *45*, 642–651. [CrossRef] [PubMed]
- 142. Kuipers, S.J.; Nieboer, A.P.; Cramm, J.M. The need for co-creation of care with multi-morbidity patients—A longitudinal perspective. *Int. J. Environ. Res. Public Health* **2020**, *17*, 3201. [CrossRef] [PubMed]
- 143. Vestjens, L.; Cramm, J.M.; Nieboer, A.P. Quality of primary care delivery and productive interactions among community-living frail older persons and their general practitioners and practice nurses. *BMC Health Serv. Res.* **2019**, *19*, 1–12. [CrossRef]
- 144. Kim, J. The effect of patient participation through physician's resources on experience and wellbeing. *Sustainability* **2018**, 10, 2102. [CrossRef]
- 145. Nguyen Hau, L.; Thuy, P.N. Customer participation to co-create value in human transformative services: A study of higher education and health care services. *Serv. Bus.* **2016**, *10*, 603–628. [CrossRef]
- 146. Osei-Frimpong, K. Patient participatory behaviours in healthcare service delivery: Self-determination theory (SDT) perspective. *J. Serv. Theory Pract.* **2017**, 27, 453–474. [CrossRef]
- 147. Hau, L.N.; Tram Anh, P.N.; Thuy, P.N. The effects of interaction behaviors of service frontliners on customer participation in the value co-creation: A study of health care service. *Serv. Bus.* **2017**, *11*, 253–277. [CrossRef]
- 148. Suárez-Álvarez, L.; Suárez-Vázquez, A.; del Río-Lanza, A.B. Companion cocreation: Improving health service encounters of the elderly. *J. Serv. Mark.* **2020**. [CrossRef]
- 149. Higgins, A.; Downes, C.; Monahan, M.; Hevey, D.; Boyd, F.; Cusack, N.; Gibbons, P. Evaluation of a co-facilitated information and learning programme for service users: The EOLAS programme. *Ir. J. Psychol. Med.* **2020**, *37*, 89–98. [CrossRef] [PubMed]

- 150. Sharma, S.; Wallace, L.M.; Kosmala-Anderson, J.; Turner, A. A process evaluation using a Self Determination Theory measure of the co-delivery of self management training by clinicians and by lay tutors. *Patient Educ. Couns.* **2013**, *90*, 38–45. [CrossRef] [PubMed]
- 151. Buckley, B.J.R.; Thijssen, D.H.J.; Murphy, R.C.; Graves, L.E.F.; Whyte, G.; Gillison, F.; Crone, D.; Wilson, P.M.; Hindley, D.; Watson, P.M. Preliminary effects and acceptability of a co-produced physical activity referral intervention. *Health Educ. J.* **2019**, *78*, 869–884. [CrossRef]
- 152. Sudhipongpracha, T.; Wongpredee, A. Demystifying Decentralization and Its Setback: Evidence From Thailand's Decentralization Reform. *Int. J. Public Adm.* **2016**, 39, 437–448. [CrossRef]
- 153. Grootaert, G.; Narayan, D.; Nyhan Jones, V.; Woolcock, M. *Measuring Social Capital: An Integrated Questionnaire*; The World Bank: Washington, DC, USA, 2004; ISBN 0-8213-5661-5.
- 154. Barrera, M.; Sandler, I.N.; Ramsay, T.B. Preliminary development of a scale of social support: Studies on college students. *Am. J. Community Psychol.* **1981**, *9*, 435–447. [CrossRef]
- 155. Centers for Disease Control and Prevention. National Survey of Children with Special Health Care Needs. Available online: https://www.cdc.gov/nchs/slaits/cshcn.htm (accessed on 13 February 2021).
- 156. Jagri, F.; Yeboah, T. Moving from rhetoric into action: Effectiveness of state–civil society synergy for healthcare delivery in rural Ghana. *GeoJournal* **2016**, *81*, 411–425. [CrossRef]
- 157. Battaglio, R.P.; Hall, J.L. Trinity Is Still My Name: Renewed Appreciation for Triangulation and Methodological Diversity in Public Administration. *Public Adm. Rev.* **2018**, *78*, 825–827. [CrossRef]
- 158. Roche, V.; Palmer, B. The Hidden Patient: Addressing the Caregiver. Am. J. Med. Sci. 2009, 337, 199–204. [CrossRef] [PubMed]
- 159. Ens, T.A.; Seneviratne, C.C.; Jones, C.; King-Shier, K.M. Factors influencing medication adherence in South Asian people with cardiac disorders: An ethnographic study. *Int. J. Nurs. Stud.* **2014**, *51*, 1472–1481. [CrossRef] [PubMed]
- 160. Layard, P.R. Cost-Benefit Analysis; Layard, R., Glaister, S., Eds.; Cambridge University Press: Cambridge, UK, 1994.