

Jun 25th, 9:00 AM

Evaluating social innovations: How creative evaluation can help articulate their values and impacts

Violet Owen

ImaginationLancaster, Lancaster University

Pinar Ceyhan

ImaginationLancaster, Lancaster University

Leon Cruickshank

ImaginationLancaster, Lancaster University

Elisavet Christou

ImaginationLancaster, Lancaster University

Follow this and additional works at: <https://dl.designresearchsociety.org/drs-conference-papers>



Part of the [Art and Design Commons](#)

Citation

Owen, V., Ceyhan, P., Cruickshank, L., and Christou, E. (2022) Evaluating social innovations: How creative evaluation can help articulate their values and impacts, in Lockton, D., Lenzi, S., Hekkert, P., Oak, A., Sádaba, J., Lloyd, P. (eds.), *DRS2022: Bilbao*, 25 June - 3 July, Bilbao, Spain. <https://doi.org/10.21606/drs.2022.607>

This Research Paper is brought to you for free and open access by the DRS Conference Proceedings at DRS Digital Library. It has been accepted for inclusion in DRS Biennial Conference Series by an authorized administrator of DRS Digital Library. For more information, please contact dl@designresearchsociety.org.

Evaluating Social Innovations (SI), how Creative Evaluation (CE) can help articulate their values and impacts

Violet Owen*, Pinar Ceyhan, Leon Cruickshank, Elisavet Christou

Imagination Lancaster, Lancaster University, UK

*corresponding e-mail: r.v.owen@lancaster.ac.uk

<https://doi.org/10.21606/drs.2022.607>

Abstract: Locally, nationally and internationally, Social Innovations (SI) are increasingly seen as a way to address the complex problems posed by society. Emphasised by both funding bodies and in UK legislation and initiatives such as Public Services (Social Value) Act 2012 and Social Impact Bonds, there is a greater urgency in evidencing the value generated by SIs. However, the process of measuring and evidencing social value is still underdeveloped. This necessitates developing evaluation approaches that are adaptive, responsive to context, and able to demonstrate value beyond financial return. This paper presents an overview of the current evaluation methods employed to capture the social value generated by SI's and examines the problems with these methods. Furthermore, it reviews Creative Evaluation (CE), a constellation of evaluation approaches, which has recently garnered renewed attention in evaluation research, and presents it as a promising avenue that could help mitigate the current issues faced when evaluating SI's.

Keywords: creative evaluation; social innovation; social impact, social value, evaluation practice

1. Introduction

Social Innovation (SI) has been increasingly viewed as vital to addressing the wide array of complex problems society poses (Murray et al., 2008). Interest in using SI's has broadened from its roots in the third sector and has become more widely acknowledged across other sectors, such as in government and in business. In the UK, many new policies and initiatives have been introduced which indicate the scaling up of SIs. The Public Services (Social Value) Act was established in 2012 and is described as “a tool to help commissioners get more value for money out of procurement;” and from 2018 central government in the UK announced that they would be explicitly evaluating “social value when awarding most major contracts” and that “government departments will be expected to report on the social impact of their major contracts” (Cabinet Office, 2021). In 2010 the first Social Impact Bond (SIB) (Mchugh et al., 2013) was awarded. SIBs are a “form of Payment by Results [... that



harness] social investment from capital markets to meet needs arising from budget cuts” (McHugh et al., 2013, p. 3). In the UK finance sector, the Big Society Capital social investment bank was established, connecting capital to social enterprises, charities and social purpose organisations (Cohen, 2012) through social impact investing. These examples are indicative of the growth SI has experienced; furthermore, the European Commission’s Single Market Act (European Commission, 2010) shows that this growth is not limited to the UK context. SI’s growing prevalence has also increased the demand for evidence of the value and impact that they generate. Evaluation plays an essential role in evidencing the value created by innovations; however, the process of measuring and evidencing social value is underdeveloped (Daykin et al., 2017). This paper argues that many evaluation approaches used to evaluate SI have been appropriated from the financial sector, and as a result they struggle to represent values generated by SIs beyond that of financial return. There are also many approaches, particularly in the UK context, which are legacy from New Public Management and financial recessions (Forss et al., 2011) and have been designed underpinned by parsimonious values (Liston-Heyes and Liu, 2021) increasing their tendency towards financial metrics. This highlights a need for designing evaluation that is adaptive, responsive to context, and able to demonstrate types of value that extend beyond value for money or profitability. This paper presents an overview of the current methods employed to capture the social value generated by SIs and through examining examples of practice, identifies the advantages and shortfalls of these approaches. This paper then discusses Creative Evaluation (CE), an area of evaluation that is currently underdeveloped but has the potential to mitigate some of the obstacles faced in establishing impact and value in SI’s.

2. Literature review

2.1 Social Innovation (SI)

SIs can exist as “specific ideas, actions, frames, models, systems, processes, services, rules and regulations as well as new organisational forms” (Nicholls et al., 2015, p. 2) aimed at meeting social needs (Milley et al., 2018), addressing intractable issues (Manzini, 2015, p. 12), and creating new social relationships or collaborations (Murray et al., 2010, p. 3), to improve social conditions, create social impact, generate social value, provide solutions to social problems (Fairbairn, 2017), create new social practices (Howaldt & Schwarz, 2010, p. 26), foster inclusion (Moulaert et al., 2013), and enhance society’s ability to act (Manzini, 2015, p. 11). SI is both an interdisciplinary and transdisciplinary practice, which transfers or borrows methodologies and examines problems whilst considering and/or utilising multiple relevant perspectives (Ziegler, 2017). SI’s transdisciplinarity allows for the inclusivity and acknowledgement of a broad spectrum of actors (Soule et al., n.d.). The primary motivation for SI is to generate social value (Murray et al., 2010, p. 3) demarcating SI from other innovation types; business, technology or economic innovation, often aim to produce a product or asset (Neumeier, 2012) for profitability (Bock, 2012; OECD & Eurostat, 2018), whereas the main beneficiary of SIs should be society (Fairbairn, 2017).

SI is often remarked on as a new or emergent trend (Szijarto et al., 2018; Tucker, 2014, p. 20), and whilst there is evidence to suggest that interest in SI has increased since the 2000's (Adams & Hess, 2010; Fairbairn, 2017; Klein, 2013, p. 9; Milley et al., 2018; NESTA, 2008; Szijarto et al., 2018), there are also many well documented examples of innovations that are focused on caring for humans and caring for the community that significantly predate this time. Examples include: microcredit such as the Grameen bank devised by Muhammad Yunus in 1976 which provided credit without the requirement of collateral (Giridharadas & Bradsher, 2006); social reforms, for example James Phillips Kay's (1832) and Edwin Chadwick's (1842) respective treaties that highlighted poor working conditions of factory workers and led to reforms (Kehler, 2008); cooperatives and credit unions like Friedrich Wilhelm Raiffeisen's credit union established in 1842 for impoverished farmers (Fairbairn, 2017); philanthropic business leaders that created model towns, such as Robert Owen's New Lanark Mill in which he created an entirely new societal structure during the early 1800's (Mulgan et al., 2007, p. 10); and model schools and homes for the destitute, like Barnardo's Ragged School, established in 1867, which gave impoverished children access to free education and accommodation (Harris & White, 2013).

Increasingly, these activities which have previously been labelled as radical innovations (Mulgan et al., 2007), social reform or socialism (Godin, 2012) are being identified as SIs. There are also examples of innovations that only became appreciated as such upon later reflection, often getting categorised as other innovation types, such as technology or business innovations, but that also had significant social impact. For example, the GNU project (St. Laurent, 2004) which is most obviously a technology innovation also generated significant social impact, as its unlimited access enabled experts and amateurs alike to freely run the software, as well as copy, distribute, modify and study it.

2.2 Evaluation in Social Innovation Contexts

As the demand for evidence has increased (Manohar et al, 2016; Mulgan 2010; Cabinet Office, 2021), so has the demand for evaluation (Vasin et al., 2017). Evaluation is often driven by funders of SI (Yee et al., 2020), to justify spending or to secure further investment; however, evaluation is also designed to aid the development of the innovation process, such as in Developmental Evaluation (Patton, 2011) or Evaluation Capacity Building (Labin et al., 2012), to understand why the innovation was or was not successful, to process qualitative data and to accurately represent positive change (Weaver & Kemp, 2017). Evaluation can also provide valuable data to enable an SI to scale up. Scaling up is often a long-term goal of SIs (Patton, 2011), as growth indicates its success, but also validation that it is an innovation. Evaluation is also conducted in order to better understand the social value created by SI, which, as explored below, is demonstrated in a variety of ways.

2.3 Social Value and Impact

Social value is often referred to as the "wider social, economic and environmental benefits that accrue from public and social investment," (Dayson, 2017, p. 395) the value of which

benefits groups or the whole of society as opposed to private gain or for profitability (Phills et al., 2008). Social value often represents the less tangible data generated by an SI that is difficult to represent through conventional means such as increase in sociability, or improved sense of contentment; these are sometimes referred to as soft issues (Black, 1995). Stakeholder groups can have very different priorities on which values should be measured and imposing or restricting how and what is to be evaluated can cause conflict, tension or mistrust to arise (Weaver & Kemp, 2017). There is often a tenuous balance between what funding bodies might ask to be reported or captured, and other values that other stakeholder groups feel are important to establish. Often there is not enough funding allocated to evaluation activities to capture a broad variety of impacts and compromises must be made. These conflicts can often leave evaluators¹ caught in an existential crisis, unsure of the evaluation's purpose.

Impact is a typical measure for assessing the success of an intervention or project. The impact generated by an innovation is often considered to be the lasting change the innovation generates to a product, process, or system. It is also only usually possible to measure impact after enough time has passed, and to be able to observe notable change has occurred. Value and impact are often discussed in tandem, value often referring to financial gain (though value is often considered differently in SIs), and impact often thought of as change generated by the intervention, which can exist from a small scale (such as change within a community) to a revolutionary scale (where whole systems and structures are transformed) (Westley & Antadze, 2010).

3. Evaluation of Social Innovations

There are many evaluation frameworks, methodologies, tool kits, and approaches that provide theories and practical methods for capturing social value or impact. Many of these approaches emphasize the return on investments generated (Nicholls, 2017; Weaver & Kemp, 2017). This section will provide an overview of some of the approaches currently used to measure the social value and social impact; as Mulgan et al (2010) note there are over 150 innovation metrics, far too many to fully document here. The examples presented in this paper are those commonly employed by programs and projects in receipt of funding allocated by government bodies. The authors will also explore how these approaches aim to capture value and impact, drawing out their advantages and shortfalls.

3.1 Social Impact Assessment and Social Return on Investment (SROI)

Increasingly popular within the non-profit world (Murray et al., 2010) are Social Impact Assessment and SROI, which are frameworks that combine principles from social accounting and cost-benefit analysis (Nicholls et al., 2009). They estimate the “direct costs of an action/intervention, the probability of it working and the likely change in future outcomes” (Weaver & Kemp, 2017, p. 9). They use a formula to represent data from non-monetary

¹ Evaluators here refers to anyone conducting an evaluation, not specifically a professional evaluator.

sources in order to make a wider variety of values comparable and reportable (Nicholls, 2017). Bosco et al., (2019) report on a SROI analysis of the *Imagine Arts Programme*, exploring the benefits generated for elderly people living in residential homes by engaging them in creative activities. They purported that engagement in creative activities can increase community inclusion, improve mental and physical health, improve cognition, decrease isolation, increase motivation, and improve the quality of life for participants; and specific to the care staff: improve their skills in caring for older people and increase their confidence in using arts interventions. Through this approach they were able to demonstrate that the benefits generated by these activities would lead to a return of 1.2 times the investment. This type of metric is useful in identifying the fiscal values generated from non-monetary sources in programmes; however, because it converts social value to economic, the outcomes it demonstrates are very one dimensional, and the emphasis is on financial return rather than the social benefits created.

3.2 Cost-Benefit Analysis and Cost-Effectiveness Analysis

Cost-Benefit Analysis (CBA) and Cost-Effectiveness Analysis (CEA) are often used in large scale program evaluations (Weaver & Kemp, 2017). They aim to quantify financially the costs and benefits of a product, program or intervention (Murray et al., 2010, p. 102). For example, a CEA was conducted to evaluate *Weight Watchers* and the *Lighten Up to a Healthy Lifestyle* weight loss program (Cobiac et al., 2010). This study compared the costs of participants attending weight loss programs, which incorporated dietary counselling and exercise advice, with the projected costs associated with caring for illnesses associated with obesity, calculating the projected saved or avoided health costs that could be accrued if participants did not improve their health. In this example, the results found that this mode of intervention was not economically beneficial when compared to the projected health gains, and the authors would not recommend this type of intervention. Here it is possible to see how CEA approaches could be potentially damaging, as the negative recommendation relies heavily on the cost-effectiveness of the intervention, paying little regard to other values that may have been generated.

3.3 Life Satisfaction Assessment, Stated Preferences and Revealed Preferences

Life Satisfaction Assessment, Stated Preferences and Revealed Preferences are all forms of monetary valuation used to understand the impact a policy, good or life event has had on a person's welfare, wellbeing or happiness (Fujiwara & Campbell, 2011). Life Satisfaction Assessment measures the amount of extra income it would take to make an equivalent degree of gain in life satisfaction. Stated Preferences measures how much a person would be willing to pay for or to forgo a good or service, and Revealed Preferences looks to uncover the perceived value of non-market goods. These approaches to evaluation can be problematic as the lens it provides is a very narrow one. The data they collect is highly commodified and often does not permit exploration into factors that may improve lives and that do not necessarily cost anything.

3.4 Quality-Adjusted Life Years Assessment and Disability Adjusted Life Years Assessment

Quality-Adjusted Life Years Assessment and Disability Adjusted Life Years Assessment are commonly used evaluation approaches in the healthcare sector. Their aim is to capture the benefits gained through improving the quality of someone's life versus the cost of the intervention (Antadze & Westley, 2012). They can provide data on the effectiveness of clinical treatments, and how they increase the quality of the remainder of a person's life (La Puma, 1992). However, their findings can be controversial. There are often serious ethical concerns with this approach, as the assessment may reveal it is not cost effective to spend money to improve the quality of a person's life (Antadze & Westley, 2012). There is also a large degree of variability in results dependant on which metrics are used which can call into question their accuracy (Marra, 2007; La Puma, 1992).

3.5 Value-Added Assessment

Value-Added Assessment is an example of evaluation that utilises numerical datasets, comparing input data with output data to judge the value generated between the input and output stages. For example, the projected grades of a student entering secondary school compared with their final attainment. This data can be used to assess how much the education provided by schools has benefitted its pupils (Şen et al., 2020). However, this type of measure is highly quantified and is unable to take into consideration individual experiences when calculating value gained. For example, a student may have had to overcome significant personal circumstances, in which the school supported that student and enabled them to stay in education, but this achievement would not necessarily be reflected in their results. In fact, their grades might suggest they have made no progress and the school has not had any positive impact on their education.

3.6 Patient Reported Outcome Measurements (PROMS)

PROMS are tools used to assess a "patient's health, quality of life, or functional status associated with health care or treatment" (Weldring & Smith, 2013, p. 62) from the patient's perspective. The data for PROMS are often collected using pre and post operative questionnaires (NHS, 2022) and is reported using statistics. They were implemented to enable patients to evaluate the care they received. However, the questionnaires used to collect patient feedback, are highly quantified and standardised. Questionnaires often tend to reflect the views of the person writing them regardless of how objective that person aims to be (Gray, 2018, p. 344; Heild, 2012), so they are an inherently problematic evaluation method. For example, respondents may select a best fit answer out of the choices provided so the data may be only partially truthful.

What can be seen from the examples above is a tendency to rely on evaluations that focus on financial return or heavily quantified results. These types of evaluation are often attractive to use. They can make it easier to compare an innovation's investment (input) with its impact (outcomes), and they can be very effective at capturing certain types of

impact, such as financial return. Quantified data can also be seen as easier to understand, and more absolute when reporting impact. SIs are often perceived to be riskier than other innovation activities (Moore et al., 2012) as they often don't have an obvious financial gain (which is another reason that using metrics that represent social value financially is appealing). This can lead to increased scrutiny or pressure from funders regarding expenditures, and lead to funders exerting control over how and what to evaluate. However, reporting social value and social impact this way can be reductive and is often ineffective in communicating social value. This can be problematic for establishing the value and impact a SI has generated; it might only demonstrate a shadow of the value created. Additionally, there is often a tendency to treat evaluation as a summative activity which can mean the SI's full impact is often not measured or not measurable. In other approaches, such as Developmental Evaluation (Patton, 2011), this has also been highlighted as an issue to address.

What is needed for evaluation practice to move forward and be better at evaluating SIs, are approaches that do not merely quantify data, but aim to construct more effective methods of reporting the value and impacts generated by SIs; and evaluation designs that promote learning and ongoing collaboration. As Preskill and Beer (2012) warn, evaluation that focuses too heavily on performance against outcomes can prevent an innovations growth, rather than supporting its development. There is also a need for evaluations that are more representative of SIs and that embody the values inherent in SIs. SIs are socially focused (Milley et al., 2018), inclusive and collaborative (Murray et al., 2010, p. 3), relationship building (Mulgan, 2007), transformative (Portales, 2019), and transdisciplinary (Penin, 2015). These characteristics need to be designed into evaluations, so they are able to successfully evaluate SIs.

4. Creative evaluation

The field of evaluation has continually evolved with developments within society. Similar to design, evaluation transformed from a professional practice to a field-specific discipline and finally into a transdiscipline. As a transdiscipline, evaluation can be designed to respond to contextual circumstances and be applied in other disciplines to address complex problems. Designing evaluation requires considering the purpose of evaluation, the characteristics of the object of evaluation and the characteristics of the problem evaluation aims to resolve (Hansen 2005), and like any design it can benefit from creativity to be effective (Patton, 1981). An under-investigated area of evaluation research is Creative Evaluation (CE), which this paper presents as a potential avenue for future research which may mitigate the current issues faced when evaluating SIs. CE was introduced by Michael Quinn Patton in 1981 in his eponymous book. It was a provocation, entreating evaluators to think more creatively about how they were practicing evaluation. Since Patton's introduction, the phrase CE has been broadly used and many definitions have emerged. More recently, it has been explored and conceptualised by Christou et al., (2021) as a constellation of evaluation approaches. By reviewing CE examples in a variety of contexts such as education, cultural industries, health,

and social work, Christou et al., (2021) identify that there is not currently a consensus on a definition of CE (it is not yet a theory, framework or model) and that the use of the term is inconsistent. They have identified that CE has been conceptualised when employing creative practices to shape the evaluation process, design evaluation activities, and develop novel evaluation methods, as well as to evaluate creativity in settings such as art, design and education. Furthermore, they have found that designing context specific evaluation solutions requires qualities often associated with creative thinking, and to do that, evaluation design employs creative methods used by other design disciplines (i.e., engagement and participation). They developed the following diagram (see Figure 1) to provide an overview of the contexts (where) and concepts (how) where CE is being used.

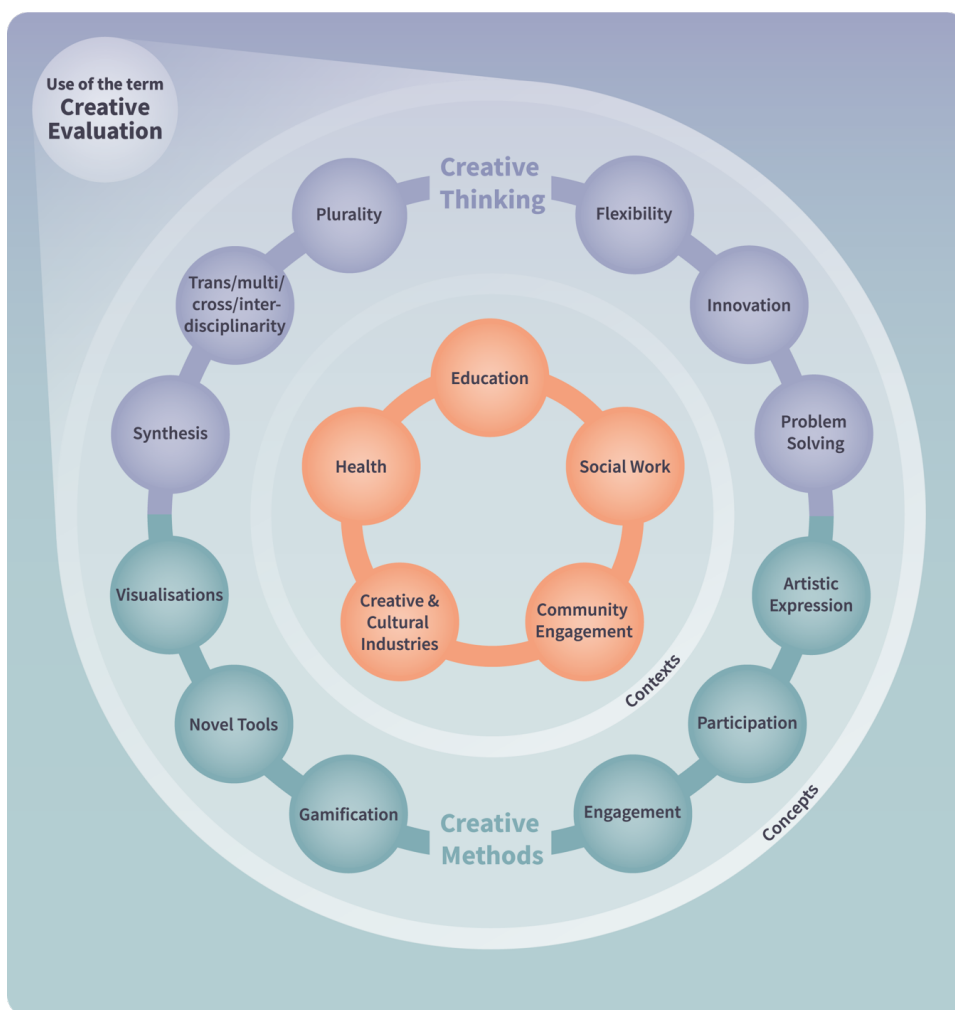


Figure 1. Diagram of uses of the term Creative Evaluation (Christou et al., 2021).

They define that,

creative thinking indicates a broader application of creativity that shapes the evaluation process and often results in the development of novel methods while creative methods indicate a more targeted use of methods that employ creative

practices in order to achieve a specific evaluation goal like increasing engagement in the evaluation process (Christou et al., 2021, p. 9-10).

However, there is much crossover between creative thinking and creative methods and often one informs the other. The following are examples drawn from a range of different contexts in which creative thinking and creative methods have been utilised in evaluations. The authors have drawn out shared themes and characteristics from these approaches to explore the benefits and opportunities they afford.

A common theme that occurs in most of the examples presented in this paper is approaches to evaluation that encourage the participation and inclusion of different groups (such as non-professionals) and inviting a broader spectrum of actors to participate in informing, designing or conducting the evaluation. Manohar et al. (2016) describe the Leapfrog Project, a collaborative project between communities, public sector partners and researchers. The evaluation tools used were co-designed during the workshops. This was done to ensure equitable involvement and to foster trust between different groups, but also to flatten hierarchies between the different actors involved. Co-designing the tools was important as the authors wanted the evaluation to promote learning and inform the future delivery of the project, as opposed to being used as an auditing tool. The authors also report that involving different stakeholders in the creation of the evaluation tools improved the visibility and understanding of results, which is beneficial to establishing impact. Edmonds et al. (2009) describe an evaluation of interactive artworks. It was conducted utilising three viewpoints, the curator, the artist and the evaluator, in a cross-disciplinary approach. This collaboration resulted in co-producing a new framework for evaluating audience experiences of interactive art. Christensen et al. (2005) present the Kiddie Focus Group, an approach that was utilised to help design a new nature centre aimed at children. As Christensen et al., (2005) identify, a key benefit of this approach was that the key users (the children) were placed at the centre of the evaluation, so the values most important to them informed the design, rather than trying to fulfil too many other needs or assuming what was most important to the children when designing.

Including different people throughout different stages of evaluation can make the process more receptive, open to adaptation, more inclusive and more human. It also creates more equitable working environments, helps to level differences in power dynamics and can grant greater agency to groups. It can also create opportunities for learning, making the evaluation reciprocal (participants also have opportunities to gain from the experience) or transformative, rather than one that just uses participants or passes judgement. Inspire to Change, an evaluation and engagement consulting firm based in Minneapolis, describe CE (and Engagement) as a framework that combines developmental evaluation, utilisation-focused evaluation, principles-focused evaluation, and arts-based methodologies. During a vlog entitled *What draws us to Creative Evaluation and Engagement*, they emphasise that CE is about putting the human experience at the centre of the evaluation, and that the purpose of evaluation should be transformational, not purely for the sake of improving programs (Inspire to Change, 2021).

Another area addressed by CE approaches is how to make the engagement of different groups in evaluation activities more possible. The Creative Communities Unit (CCU) at Staffordshire University identify that utilising creative tools and techniques can make an evaluation more accessible to a wider range of people, thus more broadly engaging. They credit this to the adaptability of the tools and techniques, and their ability to be tailored to suit the participants, subject and contexts of the evaluation. Heild (2012) describes in the Creative Evaluation Toolkit, how creative data collection methods such as using songs, poetry, scrapbooking, drawing, acting, or sculpture can be used to “enable participants to engage in understanding and share their feelings and opinions in a reflective manner” (p. 4). Stuart et al. (2015) in their book *Evaluation Practice for Projects with Young People: A Guide to Creative Research* outline a range of tools employed to help young people (and practitioners and organisations) to engage in evaluation. These include both physical tools (i.e., shields, line outs, journey maps, Lego and creative canvases) and digital tools (such as, poems, mask making, dream boards, puppetry and blob people), which are underpinned by Participatory Evaluation methodology. They believe that utilising creative tools makes evaluation more accessible to young people (as opposed to methods that might be beyond the knowledge or interest of younger participants). They also highlight in their conclusions that evaluation should be meaning-making as opposed to box-ticking, and that learning for both practitioner and participants should take place.

These examples perhaps demonstrate that creative methods can offer a higher level of engagement to a wider audience than traditional evaluation techniques. These approaches also indicate a change in the role and purpose of evaluation, from evaluation done to or about a person/group/subject to evaluation that is done with or for a person/group/subject. In many ways this echoes the development of design practices as described by Sanders and Stappers (2008) from the designing of products, to purposeful designing. These examples also exemplify the transformation of the participant, from a passive role to a stakeholder with agency. This is achieved by utilising participatory or collaborative practices in the evaluation and by making evaluations a more open process in which different actors have opportunities to participate.

Very evident in many of the examples explored here is the high degree of trans/interdisciplinarity of these approaches, many of which are at the intersection of design, art and evaluation. This is an important emergent characteristic of CE approaches, as it means they are able to draw from and build on good practice across a range of sectors and practices. It also affords a wider variety of tools that can be utilised for evaluation.

Marentakis et al. (2017) use the term CE to describe an evaluation approach at the intersection between art and Human Computer Interaction (HCI), which aims to employ interdisciplinary methods to evaluate. The authors felt the methods routinely used to capture HCI were inadequate, so looked beyond what was usually employed. Artists were invited to produce videos and artworks as evaluation, and later a written commentary to demonstrate their response to viewing interactive artworks, which were then displayed together as part of the exhibition. They felt the responses more accurately represented the

subject of the evaluation, but also, they were fully utilising artists' expertise and acumen. Approaching this evaluation as an interdisciplinary process improved the quality of the evaluation outcomes and more effectively utilised the artist evaluators.

Although there is not a common definition of CE, what can be seen from examining the examples presented here is shared characteristics surfacing in response to emergent needs in evaluation practice, of which parallels can be drawn between evaluation and design practice, particularly to co-design, participatory design and collaborative design.

4.1 Emergent Characteristics of CE

- Inclusivity: aiming to increase participation of different stakeholder groups and address any imbalances in power dynamics.
- Trans/interdisciplinarity: utilising tools, methods, and methodologies across sectors/philosophies.
- Human-centric: putting people at the core of the evaluation, discovering, and focusing on the most important values for them.
- Receptiveness: reacting to improve dissemination, encouraging different voices.
- Adaptability: generating evaluation approaches using emergent information, not pre-determining methods, and being flexible to different opinions.
- Novelty/Creativity: utilising different approaches to capture data.
- Reciprocity: evaluation being viewed as a reciprocal process (evaluation *for/with* rather than evaluation *at/about*).
- Learning: evaluations provide opportunities for learning.

Most, if not all, of the examples explored in this paper demonstrate an increased want for evaluations that centre around human experience, and that are contextually relevant and inclusive. They often seek to diminish differences in power structures to increase the breadth of participation. They also demonstrate trans/interdisciplinarity. Evaluation has predominantly evolved from sectors not traditionally associated with creativity (i.e., Finance, Business), and many of the CE tools explored in this paper are derived from the Arts or Design. This shows how evaluation is being developed across disciplines. Many of the approaches demonstrate that they are also novel, receptive and adaptative as they are being developed in situ by stakeholders. Often evaluations are conducted aiming to establish success based on predetermined goals and objectives as decided by funding bodies (Yee et al., 2020), rather than being informed by those involved in the social innovation. This can be a great barrier to evaluation and prevent the exploration of emergent impacts. However, many of the examples explored in this paper focus on discovering and representing the values that are core to what is being evaluated (such as explored by Christensen et al., 2005).

5. Discussion

The inherent complexity of SIs makes their measurement problematic (Murray et al., 2010, p. 6). SIs are often made up of complicated systems involving multiple actors and groups, conceptualisations, and ideas; selecting what to measure and how to evaluate their intricacies is difficult. It is understandable why approaches that quantify this complexity are attractive to use. However, these methods which emphasise economic models of cause and effect are not always translatable, supportive, nor appropriate for measuring SIs (Nicholls, 2009, p. 758; Weaver & Kemp, 2017, p. 9). They may not accurately represent the context or complexity of the SI (Dayson, 2017; McHugh et al, 2013), and the full spectrum of social value generated often cannot be captured through measuring financial return or consumer satisfaction (Antadze & Westley, 2012, p. 134; Kaplan & Grossman, 2010). Trying to evaluate SIs in a simple or mechanistic way can harm the innovation process and prevent it from flourishing (Milley et al., 2018, p. 241; Preskill & Beer, 2012). It may also provide results that are inaccurate or lacking context (Nicholls, 2009).

There is also the ever-tenuous relationship between different stakeholder groups, who may have converse opinions to what should be evaluated, which can cause issues of power to arise (TESPSIE, 2014). There is often a balancing act between parsimonious accountability demanded by funders (Dayson, 2017; Chelimsky, 1997) versus demonstrating the other values the innovation is generating (Milley et al., 2018). Demonstrating social value is challenging. It does not always provide evidence of the type of value recognised by funders, which can prevent social innovators from finding funding, limiting their growth, commencement, or continuation (Antadze & Westley, 2012; Bugg-Levine et al., 2012). The necessity of securing funding can also lead to poor choices for evaluating SIs. This can force evaluators and innovators to choose evaluation methods that demonstrate the credibility of the intervention over the impact it is producing (TESPSIE, 2014), and choose evaluation that is focused on outcomes rather than approaches that could improve or develop the SI. There is significant diversity within the SI landscape (Weaver & Kemp, 2017). To effectively evaluate SI, evaluation practice also needs to be similarly diverse. SI creates unique and unexpected environments for evaluation; evaluation designs are needed that not only provides evidence for financiers, but can also represent other outcomes, reflect multiple stakeholder perspectives (Dayson, 2017), and is able to effectively represent context (Kaderabkova & Moghadam Saman, 2013).

One promising way forward to resolving the systemic issues rooted in designing evaluation for SIs is CE approaches. Although CE has not yet been developed into a cohesive framework, model, methodology or theory; the characteristics identified in this paper indicate it has the potential to meet the emergent needs that SI poses. CE represents approaches that are adaptive, flexible, context rich and transdisciplinary, and that are “continually evolving in response to developments in society, economy, culture and politics” (Christou et al., 2021). CE approaches encourage collaboration, co-design and cocreation in evaluation practice which has the potential to alleviate issues regarding conflict, tension, and mistrust. There are also interesting transdisciplinary practices that have emerged

through CE approaches, such as utilising co-design to generate evaluation tools as described by Manohar et al., (2016), which increase participation, alleviate tensions and difference in power between stakeholders and improve dissemination of results. Through examination of the examples in this paper, it is possible to see how CE is working towards developing the same principles that underpin SI, which is a crucial consideration for evaluation practice. More research and practice are needed to further establish CE as an approach, framework or methodology and to improve its efficacy in evaluating the social value and impact of SIs; however, it has the potential to make valuable contributions to the evaluation of SIs in the future.

6. Conclusion

There is a growing body of evaluation research that similarly identifies the emergent needs for evaluating SIs addressed in this paper, and that suggest new directions for evaluation research. They also further articulate the issues being faced by evaluators of SIs, many of whom have suggested frameworks and principles to address these. For example, Schwandt and Gates (2021) identify the changing role of evaluation and evaluators from external entities that pass judgement to thoughtful partners working with enterprises. The Centre for Cultural Value based in Leeds University has recently co-created a set of evaluation principles to be adopted within the cultural sector (Centre for Cultural Value, 2021).

Pioneers such as J. Bradley Cousins, Elisabeth Whitmore, Michael Quinn Patton and David Fetterman who developed frameworks for Participatory Evaluation (Cousins and Whitmore, 1998), Developmental Evaluation (Patton, 2011) and Empowerment Evaluation (Fetterman, 2019) have undeniably shaped evaluation practice. Additionally, Joyce Yee, Yoko Akama and Khemmiga Teerapong developed the Designing Social Innovation Evaluation Framework which promotes a community led approach to evaluation (Yee et al, 2020). These are all indicators of evaluation's evolution.

The evaluation landscape is also becoming more openly interdisciplinary and there is much more research being published at the crossroads of design, the arts and evaluation, and the benefits of these collaborations explored. Despite this growing body of research, there is much space and scope for evaluation practice to continue to grow. The current methods of evaluation that are routinely used to evaluate SIs are in many ways inadequate. Too much emphasis is placed on profitability and return on investment, which provides too narrow a view to fully capture the range of different values generated by SIs (Potter, 2014). There is a need for evaluation that addresses these barriers and enables SIs to flourish. Evaluation designs that are able to encompass and promote the core values underpinning SIs are needed. Approaches such as CE have the potential to mitigate challenges faced by social innovators and have great potential for future development into cohesive methods, tools, frameworks or principles.

Acknowledgements: This research was carried out as part of the Beyond Imagination project at ImaginationLancaster funded by Research England Expanding, Excellence in England (E3).

7. References

- Adams, D., & Hess, M. (2010). Social Innovation and Why it Has Policy Significance. *The Economic and Labour Relations Review*, 21(2), 139-155. doi:10.1177/103530461002100209
- Antadze, N., & Westley, F. R. (2012). Impact Metrics for Social Innovation: Barriers or Bridges to Radical Change? *Journal of Social Entrepreneurship*, 3(2), 133-150. doi:10.1080/19420676.2012.726005
- Black, S. (1995). An empirical model for total quality management. *Total Quality Management*, 6(2), 149-164. doi:10.1080/09544129550035495
- Bock, B. B. (2012). Social innovation and sustainability; how to disentangle the buzzword and its application in the field of agriculture and rural development. *Studies in Agricultural Economics*, 114, 57-63.
- Bosco, A., Schneider, J., & Broome, E. (2019). The social value of the arts for care home residents in England: A Social Return on Investment (SROI) analysis of the Imagine Arts programme. *Maturitas*, 124, 15-24. doi:10.1016/j.maturitas.2019.02.005
- Bugg-Levine, A., Kogut, B., & Kulatilaka, N. (2012). A New Approach to Funding Social Enterprises. *Harvard Business Review*, 90(1/2), 118-123. Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=69948711&site=ehost-live&authtype=ip,shib&user=s1523151>
- Cabinet Office, (2021). Social Value Act: information and resources. Retrieved from <https://www.gov.uk/government/publications/social-value-act-information-and-resources/social-value-act-information-and-resources>
- Centre for Cultural Value. (2021). Evaluation Principles. University of Leeds. <https://www.culturalvalue.org.uk/our-work/evaluation/evaluation-principles/#beneficial>
- Chadwick, E. (1842). Report on the Sanitary Condition of the Labouring Population of Great Retrieved from Edinburgh:
- Chelimsky, E. (1997). The Coming Transformations in Evaluation. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st Century: A Handbook*. doi:10.4135/9781483348896
- Christensen, L., Nielsen, J. E., Rogers, C. M., & Volkov, B. (2005). Creative data collection in nonformal settings. *New Directions for Evaluation*, 2005(108), 73-79. doi:10.1002/ev.172
- Christou, E., Owen, V., & Ceyhan, P. (2021). Concepts and Contexts of Creative Evaluation Approaches. *Journal of MultiDisciplinary Evaluation*, 17(40), 1-15.
- Cobiac, L., Vos, T., & Veerman, L. (2010). Cost-effectiveness of Weight Watchers and the Lighten Up to a Healthy Lifestyle program. *Australian and New Zealand Journal of Public Health*, 34(3), 240-247. doi:10.1111/j.1753-6405.2010.00520.x
- Cohen, S. R. (2012). Big Society Capital Marks a Paradigm Shift. *Stanford Social Innovation Review*, 10(3), 21-22. <https://doi.org/10.48558/M8KZ-TB48>
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. *New Directions for Evaluation*, 1998(80), 5-23. <https://doi.org/10.1002/ev.1114>
- Creative Communities Unit. (2021). Creative and Participatory Evaluation. Staffordshire University. <http://blogs.staffs.ac.uk/ccu/current-research-and-projects/creative-and-participatory-evaluation/>
- Daykin, N., Gray, K., Mccree, M., & Willis, J. (2017). Creative and credible evaluation for arts, health and well-being: opportunities and challenges of co-production. *Arts & Health*, 9(2), 123-138. doi:10.1080/17533015.2016.1206948

- Dayson, C. (2017). Evaluating social innovations and their contribution to social value: the benefits of a 'blended value' approach. *Policy & Politics*, 45(3), 395-411. doi:10.1332/030557316x14564838832035
- Edmonds, E., Bilda, Z., & Muller, L. (2009). Artist, evaluator and curator: three viewpoints on interactive art, evaluation and audience experience. *Digital Creativity*, 20(3), 141-151. <https://doi.org/10.1080/14626260903083579>
- European Commission. (2010). *Your Single Market? Single Market Act for a highly competitive social market economy*. Belgium: European Commission
- European Commission. (2014). *Proposed Approaches to Social Impact Measurement in European Commission legislation and in practice relating to EuSEFs and the EaSI: GECES Sub-group on Impact Measurement 2014*. Luxembourg: European Commission
- Fairbairn, B. (2017). RAIFFEISEN AS SOCIAL INNOVATOR. *Annals of Public and Cooperative Economics*, 88(3), 425-448. doi:10.1111/apce.12157
- Fetterman, D. (2019). Empowerment evaluation: a stakeholder involvement approach. *Health Promotion Journal of Australia*, 30(2), 137-142. <https://doi.org/10.1002/hpja.243>
- Forss, K., Marra, M. and Schwartz, R. (2011). *Evaluating the complex*. New Brunswick, NJ: Transaction Publishers.
- Fujiwara, D., & Campbell, R. (2011). Valuation Techniques for Social Cost-Benefit Analysis: Stated Preference, Revealed Preference and Subjective Well-Being Approaches. A Discussion of the Current Issues. HM Treasury Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/209107/greenbook_valuationtechniques.pdf
- Giridharadas, A., & Bradsher, K. (2006). Microloan Pioneer and His Bank Win Nobel Peace Prize. *The New York Times*. Retrieved from <https://www.nytimes.com/2006/10/13/business/14nobelcnd.html>
- Godin, B. (2012). *Social Innovation: Utopias of Innovation from c.1830 to the Present*. Retrieved from http://www.csiic.ca/PDF/SocialInnovation_2012.pdf
- Gray, D. E. (2018). *Doing Research in the Real World* (4th ed.). Sage Publications.
- Harris, J., & White, V. (Eds.). (2013) *A Dictionary of Social Work and Social Care*. Oxford University Press.
- Hansen, H. F. (2005). Choosing evaluation models: A discussion on evaluation design. *Evaluation*, 11(4), 447-462.
- Heild, A. (2012). *Creative Evaluation Toolkit*. <https://www.artworkscreative.org.uk/wp-content/uploads/2015/11/Creative-Evaluation-Toolkit.pdf>
- Howaldt, J., & Schwarz, M. (2010). *Social Innovation: Concepts, Research Fields and International Trends* (Vol. 5). Dortmund: Aachen.
- Inspire to Change (2021). I2C Vlog 001 What draws us to Creative Evaluation and Engagement <https://www.youtube.com/watch?v=XuBIM6zIWOM>
- Kaderabkova, A., & Moghadam Saman, S. (2013). Evaluation of social innovations – Their characteristics and impacts, cross county comparisons and implications for policy support. Paper presented at the Social Frontiers: The next edge of social innovation research, London.
- Kaplan, R. S., & Grossman, A. S. (2010). The Emerging Capital Market for Nonprofits. *Havard Business Review*. Retrieved from <https://hbr.org/2010/10/the-emerging-capital-market-for-nonprofits>
- Kay, J. P. (1832). *The Moral and Physical Condition of the Working Classes Employed in the Cotton Manufacture in Manchester*. London: Frank Cass.

- Kehler, G. (2008). Gothic Pedagogy and Victorian Reform Treatises. *Victorian Studies*, 50(3), 437-456. doi:10.2979/vic.2008.50.3.437
- Klein, J.-L. (2013). Introduction: social innovation at the crossroads between science, economy and society. In F. Moulaert, D. MacCallum, A. Mehmood, & A. Hamdouch (Eds.), *The International Handbook on Social Innovation* (pp. 9-24). Cheltenham: Edward Elgar.
- Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A Research Synthesis of the Evaluation Capacity Building Literature. *American Journal of Evaluation*, 33(3), 307-338. doi:10.1177/1098214011434608
- la Puma, J. (1992). Quality-Adjusted Life Years: Ethical Implications and the Oregon Plan. *Issues in Law and Medicine*(4).
- Liston-Heyes, C., & Liu, G. (2021). To measure or not to measure? An empirical investigation of social impact measurement in UK social enterprises. *Public Management Review*, 23(5), 687-709. <https://doi.org/10.1080/14719037.2020.1865435>.
- Manohar, A. K., Smith, M., & Calvo, M. (2016). Capturing the “How”: Showing the value of codesign through creative evaluation Paper presented at the DRS 2016, Brighton.
- Manzini, E. (2015). *Design, When Everybody Designs: An Introduction to Design for Social Innovation*. Cambridge, Massachusetts: The MIT Press.
- Marra, C. A., Marion, S. A., Guh, D. P., Najafzadeh, M., Wolfe, F., Esdaile, J. M., Clarke, A. E., Gignac, M. A., & Anis, A. H. (2007). Not all “quality-adjusted life years” are equal. *Journal of Clinical Epidemiology*, 60(6), 616-624. <https://doi.org/https://doi.org/10.1016/j.jclinepi.2006.09.006>
- Marentakis, G., Pirro, D., & Weger, M. (2017). *Creative Evaluation. Designing Interactive Systems*, Edinburgh.
- McHugh, N., Sinclair, S., Roy, M., Huckfield, L., & Donaldson, C. (2013). Social impact bonds: a wolf in sheep’s clothing? *Journal of Poverty and Social Justice*, 21(3), 247-257.
- Milley, P., Szijarto, B., Svensson, K., & Cousins, J. B. (2018). The evaluation of social innovation: A review and integration of the current empirical knowledge base. *Evaluation*, 24(2), 237-258. doi:10.1177/1356389018763242
- Moore, M.-L., Westley, F. R., & Nicholls, A. (2012). The Social Finance and Social Innovation Nexus. *Journal of Social Entrepreneurship*, 3(2), 115-132. <https://doi.org/10.1080/19420676.2012.725824>
- Moulaert, F., MacCallum, D., & Hillier, J. (2013). Social innovation: intuition, precept, concept, theory and practice. In F. Moulaert, D. MacCallum, A. Mehmood, & A. Hamdouch (Eds.), *The International Handbook on Social Innovation: Collective Action, Social Learning and Transdisciplinary Research*: Cheltenham, UK : Edward Elgar, [2013].
- Mulgan, G. (2010). Measuring Social Value. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/measuring_social_value
- Mulgan, G., Tucker, S., Ali, R., & Sanders, B. (2007). *Social Innovation: What it is, why it matters, how it can be accelerated*. Retrieved from The Basingstoke Press:
- Murray, R., Caulier-Grice, J., & Mulgan, G. (2010). *The Open Book of Social Innovation: The Young Foundation/NESTA*.
- Murray, R., Mulgan, G., & Caulier-Grice, J. (2008). *How to innovate: the tools for social innovation*. Retrieved from <https://youngfoundation.org/wp-content/uploads/2012/10/How-to-innovate-the-tools-for-social-innovation.pdf>
- NESTA. (2008). *New approaches to transforming public services*. Retrieved from London: https://media.nesta.org.uk/documents/social_innovation_policy_brief_cover.pdf

- Neumeier, S. (2012). Why do Social Innovations in Rural Development Matter and Should They be Considered More Seriously in Rural Development Research? - Proposal for a Stronger Focus on Social Innovations in Rural Development Research. *Sociologia Ruralis*, 52(1), 48-69. doi:10.1111/j.1467-9523.2011.00553.x
- Nicholls, A. (2009). 'We do good things, don't we?': 'Blended Value Accounting' in social entrepreneurship. *Accounting, Organizations and Society*, 34(6-7), 755-769. doi:10.1016/j.aos.2009.04.008
- Nicholls, A., Simon, J., & Gabriel, M. (2015). Introduction: Dimensions of Social Innovation. In *New Frontiers in Social Innovation Research* (pp. 1-26): Palgrave Macmillan UK.
- Nicholls, J. (2017). Social return on investment—Development and convergence. *Evaluation and Program Planning*, 64, 127-135. doi:10.1016/j.evalprogplan.2016.11.011
- NHS. (2022). Finalised Patient Reported Outcome Measures (PROMs) in England, for Hip and Knee Replacement Procedures (April 2020 to March 2021). NHS. <https://digital.nhs.uk/data-and-information/publications/statistical/patient-reported-outcome-measures-proms/finalised-hip-and-knee-replacement-procedures-april-2020-to-march-2021/introduction>
- OECD & Eurostat (2018), Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation, 4th Edition, The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing, Paris/Eurostat, Luxembourg, <https://doi.org/10.1787/9789264304604-en>.
- Patton, M. Q. (1981). *Creative Evaluation*. Beverly Hills ; London: Sage.
- Patton, M. Q. (2011). *Developmental Evaluation : Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.
- Penin, L., Staszowski, E., & Brown, S. (2015). Teaching the Next Generation of Transdisciplinary Thinkers and Practitioners of Design-Based Public and Social Innovation. *Design and Culture*, 7(3), 441-450. doi:10.1080/17547075.2015.1105564
- Phills, J. A., Deiglmeier, K., & Miller, D. T. (2008). Rediscovering Social Innovation. *Stanford Social Innovation Review*, 36-43.
- Portales, L. (2019). Social Impact in Social Innovations: Definition, Design, and Evaluation. In *Social Innovation and Social Entrepreneurship* (pp. 161-176). Palgrave Macmillian. https://doi.org/https://doi.org/10.1007/978-3-030-13456-3_11
- Potter, D. (2014). Why the Future of Social Innovation Is Open. *Stanford Social Innovation Review*. Retrieved from https://ssir.org/articles/entry/why_the_future_of_social_innovation_is_open
- Potter, D. (2014). Why the Future of Social Innovation Is Open. *Stanford Social Innovation Review*. https://ssir.org/articles/entry/why_the_future_of_social_innovation_is_open
- Preskill, H., & Beer, T. (2012). *Evaluating Social Innovation*. FSG.
- Public Services (Social Value) Act 2012. (2012). Retrieved from <https://www.legislation.gov.uk/ukpga/2012/3/enacted>
- Sanders, E. B.-N., & Stappers, P. J. (2008). Co-creation and the new landscapes of design. *CoDesign*, 4(1), 5-18. <https://doi.org/10.1080/15710880701875068>
- Schwandt, T., & Gates, E. F. (2021). *Evaluating and Valuing in Social Research*. The Guildford Press.
- Şen, S., Yıldırım, İ., & Karacabey, M. F. (2020). The Implementation of the Value-Added Assessment to Determine School Effectiveness. *Education and Science*, 45(202), 79-91.
- Soule, S. A., Malhotra, N., & Clavier, B. (n.d.) *Defining Social Innovation*. Retrieved from <https://www.gsb.stanford.edu/faculty-research/centers-initiatives/csi/defining-social-innovation>
- St. Laurent, A. M. (2004). *Understanding open source and free software licensing* (First edition. ed.).

- Stuart, K., Maynard, L., & Rouncefield, C. (2015). *Evaluation Practice for Projects with Young People: A Guide to Creative Research*. doi:10.4135/9781473917811
- Szijarto, B., Milley, P., Svensson, K., & Cousins, J. B. (2018). On the evaluation of social innovations and social enterprises: Recognizing and integrating two solitudes in the empirical knowledge base. *Evaluation and Program Planning*, 66, 20-32. doi:10.1016/j.evalprogplan.2017.08.010
- TEPSIE. (2014). 'Social Innovation Theory and Research: A Summary of the Findings from TEPSIE.' A deliverable of the project: "The theoretical, empirical and policy foundations for building social innovation in Europe. Retrieved from Brussels:
- Tucker, S. (2014). *Social Innovation for Public Service Excellence*. Retrieved from Singapore:
- Vasin, S. M., Gamidullaeva, L. A., & Rostovskaya, T. K. (2017). The Challenge of Social Innovation: Approaches and Key Mechanisms of Development. *European Research Studies Journal*, XX(2B), 25-45.
- Weaver, P. M., & Kemp, R. (2017). A Review of Evaluation Methods Relevant for Social Innovation with Suggestions for their Use and Development. Retrieved from
- Weldring, T., & Smith, S. M. S. (2013). Article Commentary: Patient-Reported Outcomes (PROs) and Patient-Reported Outcome Measures (PROMs). *Health Services Insights*, 6, HSI.S11093. <https://doi.org/10.4137/hsi.s11093>
- Westley, F., & Antadze, N. (2010). Making a Difference: Strategies for Scaling Social Innovation for Greater Impact. *The Innovation Journal*, 15(2), 2-19.
- Yee, J., Akama, Y., & Teerapong, K. (2020). Being community and culturally-led: Tensions and pluralities in evaluating social innovation. Paper presented at the Service Design and Innovation Conference, Melbourne, Australia.
- Ziegler, R. (2017). Social innovation as a collaborative concept. *Innovation: The European Journal of Social Science Research*, 30(4), 388-405. doi:10.1080/13511610.2017.1348935

About the Authors:

Violet Owen is a PhD Design student from ImaginationLancaster at Lancaster University, England. Her research interests include Creative Evaluation, collaborative practices, and social value and impact.

Pinar Ceyhan is an International Lecturer in Design at Lancaster University, UK. Her research interests include understanding the role of experience design in cognition and meaning-making and the methods and processes for embedding evaluation that inform design and evidence value and impact.

Leon Cruickshank is Director of Research for ImaginationLancaster and a Professor of Design and Creative Exchange at Lancaster University, UK. His research focuses on the development of open and co-design processes especially in the public sector. He leads the £13.2 million Beyond Imagination project.

Elisavet Christou is a Post-doctoral Research Associate in Evaluation at Lancaster University, UK. She is a transdisciplinary researcher and works within and across evaluation, design, computing, digital arts and media.