Author Accepted Manuscript

Akama, Y., Chen, K. L., Imanishi, H., Kikuchi, Y., Kushinsky, S., Teasley, S., … Yee, J. (2026). Reflecting and transforming together through evaluative practice in social innovation. *The Design Journal,* 1–21. https://doi.org/10.1080/14606925.2026.2726360

Reflecting and transforming together through evaluative practice in social innovation

Yoko Akama[a], Ko-Le Chen[b], Hirotake Amanashi[c], Yuko Kikuchi[d], Sarah Kushinsky[a], Sarah Teasley[a], Khemminga Teerapong[e], Joyce Yee[b]

[a]School of Design, RMIT University, Narrm, Kulin Nation (Melbourne, Australia); [b]School of Design, Northumbria University, Newcastle upon Tyne, UK; [c]Department, University, City, Country Country; [d]VARI NALA, Victoria and Albert Museum, London, UK; [e]Faculty of Fine and Applied Art, Bangkok University, Bangkok, Thailand

Evaluating the social impact of co-designing is important yet complex in the messy, nuanced, and intersectional context of social change. This paper raises concerns with often-conventional approaches in social innovation and service design that prioritise objectivity, uniformity and predetermination. These concerns magnify when Global North frameworks are applied in Global South without consideration for knowledges, including frameworks, as situated. The authors argue for diversifying evaluative practices in social innovation and service design, drawing on their reflexive learnings through a transcultural peer mentoring program for and with women in creative and social innovation sectors in the Asia Pacific. The narrative traces how the team began with an evaluative 'framework', and through its iteration, catalysed transformation in their practices. Rather than detaching the framework for reporting in service design, we discuss gradual shifts as reflexive *frameworking,* as ongoing practices of adaptation in participatory co-designing.

Keywords: service design, social impact evaluation, transcultural, feminism, participatory co-design

Introduction: the double-edge sword of evaluation

Evaluating the impact of design interventions is essential for understanding their role in social transformation, yet what is often evaluated in service design (SD) remains contested and nascent (Foglieni, Villari and Sleeswijk 2018). While some practitioners and researchers recognise the importance of evaluation and the need for a broader suite of approaches for diverse contexts, evaluation efforts remain limited to what are objectively observable and measurable (Bonilla Berrocal, De Rosa, and Auricchio 2025).

In this paper, we argue that adherence to positivist knowledge and methodologies works against maturing evaluative practices when designing with communities to deliver socially-focused outcomes. Quantitative, financial statistics and user-engagement metrics are established, common methods (Mihci 2019), often driven by funders' agendas (Kah and Akenroye 2020; Krasnopolskaya and Korneeva 2020). Communities perceive such evaluative requirements as a shackle, due to such methods being ill-fitting for longer time-scales and the multifaceted nature of change (Brandsen et al. 2016; Cünha, Alves, and Araujo 2022; Weaver and Kemp 2017). In all, pre-defining the methods, processes and outcomes of evaluation can reinforce disempowering hierarchies, disregard local priorities and practices (Groves and Hinton 2004; Lamping et al. 2024) and make it difficult to evaluate how the intended social innovation might benefit communities. When funders and institutions also commission designers and researchers to evaluate their work, this requires vigilance of our own positionalities and the ways we build our capacities in evaluative practice.

As design researchers with strong connections to Australasia, committed to social justice and community participation, we felt the need to expand and explore evaluative practices in social innovation. This paper draws insight from our collaboration for over six years.

We begin with a summary of how our team (DESIAP) co-designed an evaluative framework in 2017 with researchers and practitioners in Southeast Asia to explain why this situatedness is an important starting point. We then share how this evaluative framework was iterated and our evaluative practice further developed later in 2021, within an 18-month transcultural mentorship program with and for women in Asia-Pacific. Our pursuit of responsive and responsible evaluative practice spans time and many kinds of designing, such as co-designing the original evaluative framework in 2017 and its subsequent iteration in 2021; co-designing the mentoring program and the participatory research approaches. To note, the paper focuses mainly on our reflexive adaptation with the framework during 2021-22, to evidence how we co-designed our evaluative practices.

Research approach

As a geographically dispersed team of eight, online documents (eg. Miroboards, fig. 3) became an asynchronous space to trace, sense-make, add questions, comments and concerns throughout the program. These formed majority of the data, and in addition, formal interviews, informal catch ups with participants, meeting notes, team recordings and workshop outputs were synthesised. Analysis was further guided by *Core Values* and *Principles* and evaluative questions in the frameworks (figs. 1 and 2), to identify and draw out the most salient insights for discussion. Together, the emergent learnings demonstrate how an ongoing, reflexive, learning-based approach to evaluation is important to enable responsiveness to contingent contexts and

foster conditions for social innovation. Due to space limitations, we dive deeply into two examples.

The frameworks we used were valuable touchpoints and visual tools for facilitating reflexive conversations and eliciting meaningful input from participants and collaborators. Contributions from them strengthened adaptation and responsiveness by shaping and transforming our practices, and theirs alike. However, rather than attributing causality in social impact to methods and frameworks, we underscore the necessity of constant adaptive practice. We argue that design frameworks can only ever be situated, despite design research's preference for generalisable approaches. As such, we have coined a term *frameworking*. A situated approach requires continually questioning and rethinking our practices, including how we can respond more mindfully to the dynamics of relationships constituted by people who are differently situated. The verb, *frameworking,* helps tether the framework to this reflexive praxis. We see this as both a political and ethical stance in evaluation, (service) design, and research.

Evaluations in (service) design

Many established frameworks are available for evaluating both service design (SD) processes and service outcomes. Throughout this paper, we question how their situatedness shapes their effectiveness and appropriateness for evaluating lived experiences and knowledge rooted in many paradigms. Foglieni (2025) distinguishes between evaluating the SD process and assessing service outcomes. The former focuses on measuring the value of SD itself, whereas the latter examines the service's value as an outcome of the design process. Models such as the *Design Value Scorecard* (Westcott et al. 2013), the *McKinsey Design Index* (Sheppard et al. 2018), and the *Service Design Maturity Model* (Corsten and Pick 2019) primarily assess the value SD contributes to innovation, often assuming fixed and bounded contexts. Outcome-oriented models

such as the *Net Promoter Score* (Reichheld and Markey 2011), the *Balanced Scorecard* (Kaplan and Norton 1992), and *Return on Experience* (PWC 2019) are designed to measure transactional services and economic return on investment, rather than capturing social value for all stakeholders.

Questions about 'who' defines 'what' as valuable, and 'how' things can be known or not are present in participatory evaluation (Bossen et al. 2016; Cousins and Whitmore, 1998; Drain et al. 2021; Mertens, 2009) but such discourse are still nascent in (service) design. Allies in participatory design (PD), for example, has emphasised the importance of systematic and explicit evaluation to enhance PD's accountability. Nguyen and colleagues (2024) analysed examples that evaluated processes, outputs, outcomes and impact. Undertaking and sharing systematic evaluation can support deeper awareness of project dynamics, contributing to improving methodologies and practices while assisting transparency and accountability. They echo similar concerns we hear in participatory evaluation, around epistemic, cultural and power differences, especially when researchers in privileged positions look for 'credibility, transferability, dependability and confirmability' as a measure that defines 'trustworthiness' (Nguyen, Jaspaert, and Harthum 2024, 137). These questions are not only political and ethical matters when evaluating but shape the ability of evaluation to provide a representative picture of impact.

One barrier to maturing evaluative practices in (service) design is an epistemic preference that is external, independent judgement distanced from the knower (Tuhiwai Smith 2021), reflected in statistical, neutral, and anonymous ways of collecting data. We are informed also by discourses in decolonising evaluation to problematise rigidity and universality of the West (Corsetti 2013; Middernacht, Lluch and de Iongh 2023). Funding organisations in the Global North that aid communities in the Global South can also magnify disempowerment through

evaluation. Predetermined categories of what funders deem as important, see as valid knowledge, and prefer as evaluative methods to produce relevant evidence are often built into funding conditions. These categories then become the measures of social impact. Such approaches, here termed Traditional Evaluation, are often formative and summative (ie, conducted at the end) and administered by an outsider-expert to maintain objectivity.

When funders' preferences do not account for community values or ways of knowing and understanding their world, we see a problem. Indigenous scholars and communities in the Global South experience discrimination because their knowledge is continually judged as inferior, unverifiable and untrustworthy (Akama et al. 2022; Hui 2023; Tuhiwai Smith 2021). Learning and embracing multiple forms of knowledge – inseparable from multiple values and worldviews – is thus an ethical and political endeavour. This underscores the rationale for developing an evaluative approach that recognise plural knowledges, experiences, and ways of being.

The first adaptation: Moving away from dominant conventions

The core team, transcultural by migration, education, upbringing and family, have built a relationship-first methodology, underpinned by trust and participation (see center circle in figs. 1 and 2). This is grounded in our bodies, kinships and living experiences. We work across sectors, cultures, languages, hierarchies, geographies and worldviews. Reflexively acknowledging heterogeneity is how we live, work and relate. Through mutual learning and supporting varied ways of working across thresholds, attending to divergent intersectional experiences and sensitivities we embrace care and plurality as assets in our collaborative work, including the 2021-23 peer mentoring program.

As design researchers committed to culturally-grounded, community-led collaboration in Australasia, expanding evaluative practices in social innovation was necessary to enact our values and participatory processes. Akama and Yee (co-founders of DESIAP) embarked upon co-designing an evaluative approach with changemakers in Southeast Asia in 2017. This work has been published (Yee, Akama and Teerapong 2020); we provide a brief background here.

Over three days in 2017, practitioners and researchers from Cambodia, Indonesia, Malaysia, Myanmar, Thailand and the Philippines shared knowledge and experiences of evaluating the impact of their work in social innovation. This initiative generated insights, which were then analysed and synthesised with literature in Developmental Evaluation (Patton 2011). This research confirmed that a participatory co-design approach, with cyclical action-oriented reflection embedded throughout, is critical when evaluating social innovation, as it enables aims and approaches to be defined together with community and adapted in real time to emerging conditions.

The 2017 research generated a propositional framework for evaluating and designing impactful social innovation. We call this the 'blue-flower framework' in this paper (see fig. 1). Guiding principles were distilled from literature and lived experiences as *Core Values* (center blue circle), and *Principles* (outer white circles). Arrows around the 'petals' indicate the cyclical and ongoing nature of this work. Text linked to the white circles, offers prompts for practicing the values or principles.

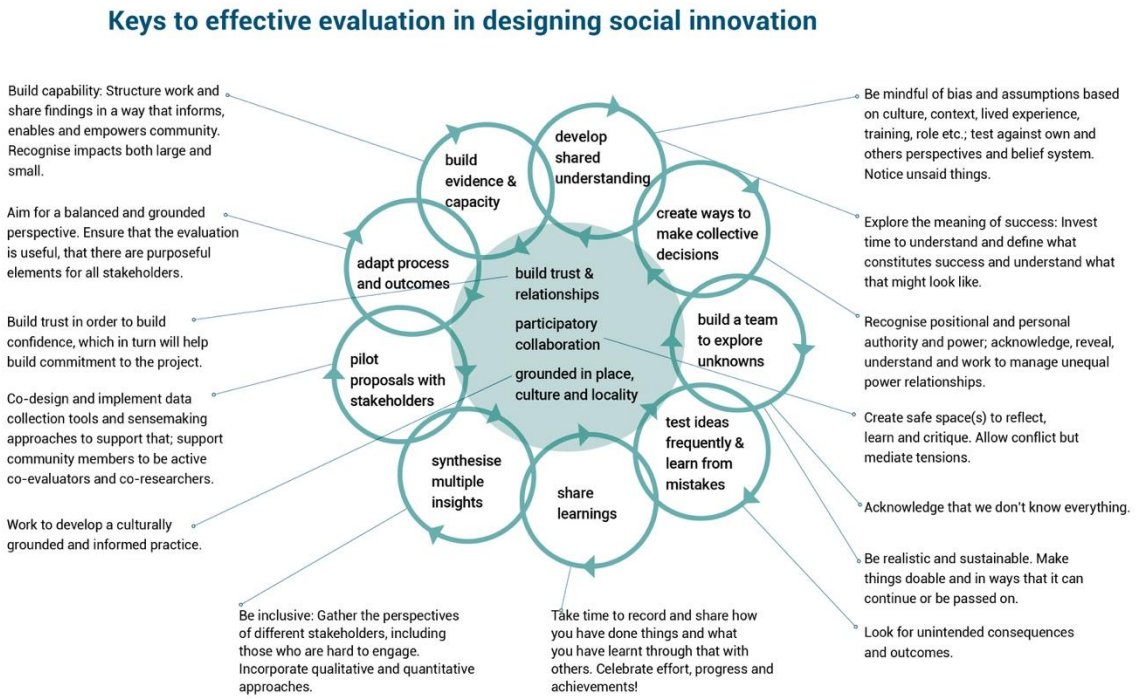


Figure 1. 'Blue-flower framework' (Akama, Yee, Hill and Tjahja 2019)

One caution: when Akama and Yee initially co-designed and published the framework (Yee, Akama and Teerapong 2020), they were concerned that due to dominant epistemes in design research, its contribution might be expected in the model's replicability and transportability (Akama and Prendiville 2013; Evans, Leah and Emina Petrović 2020). Naming it as a 'framework' is a risk to efface the dynamic collaborative practice and diverse ways of knowing. We revisit these conundrums in the discussion. The syncretic verb *frameworking* reflects this crucial point in emphasising how the framework became embodied, through adapting and responding to emerging situations.

The second adaptation: iterating the framework with trusted advisors

We iterated the ‘blue-flower framework’ (fig. 1) during an 18-month peer-mentoring program (2022-23), which grew out of a 4-month pilot in 2021 (Yee, Teerapong and Akama 2023). The longer program was conceived in response to our observation that women in East and South-East Asia have few opportunities for culturally-grounded mentoring, and to address barriers such as age and gender discrimination. The transcultural mentoring program itself developed through co-design, evaluation and reflection with the mentors on the pilot program.

The program had 37 participants, including 21 mentors and mentees, six Advisory Board (AB) members and two guest mentors. The eight members of the core program team are all co-authors of this paper. All variously work in creative, social innovation, education and research sectors in Malaysia, Taiwan, Japan, Thailand, Hong Kong SAR, Australia, Myanmar, UK, Mexico and Kenya. Some are neurodiverse and/or have physical disabilities (for more see Yee, Teerapong and Akama 2023).

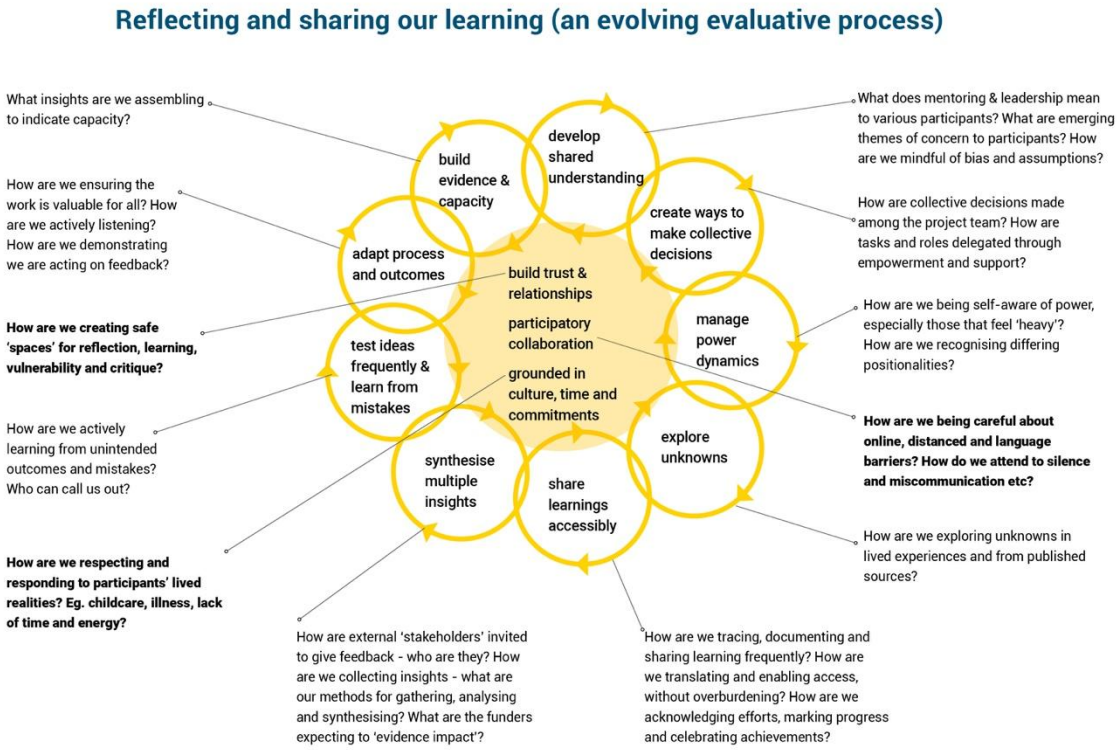


Figure 2. 'Yellow-flower framework' (Akama et al. 2022b)

Prior to the commencement of the mentoring program, Akama and Yee began iterating the blue flower framework in discussion with mentors from the 4-month pilot. Some mentors and Advisory Board (AB) members participated in the pilot and/or were past, or regular collaborators with DESIAP. Furthermore, some were acquaintances, colleagues and partners of the participants of the 2017 event, that led to co-designing the blue-flower framework. Positioning this process in the broader arc of DESIAP collaboration is important. Working with thick relationship is characteristic of DESIAP and critical for inviting frank and trusted advice. Going through this structured process helped adapt the *Core Values* (central circle) and *Principles* (outer circles). Majority remained relevant, with minor changes. Their suggestions asked for

considerations around safe participation, ways to reframe success and failure, and to anticipate and respond to disengagement. Their feedback helped refine words, concepts and ideas in early drafts for accessibility and clarity, by simplifying questions and phrasing. All together, examples of practices (eg. Principles in fig.1) were translated into context-specific questions (fig. 2) and further synthesised with an anticipated theory of change. We began to call this update, the 'yellow-flower framework' to again, share with mentors and AB members. A Miroboard (fig. 3) traced the evolution of the framework to document ongoing reflections, input and discussions.

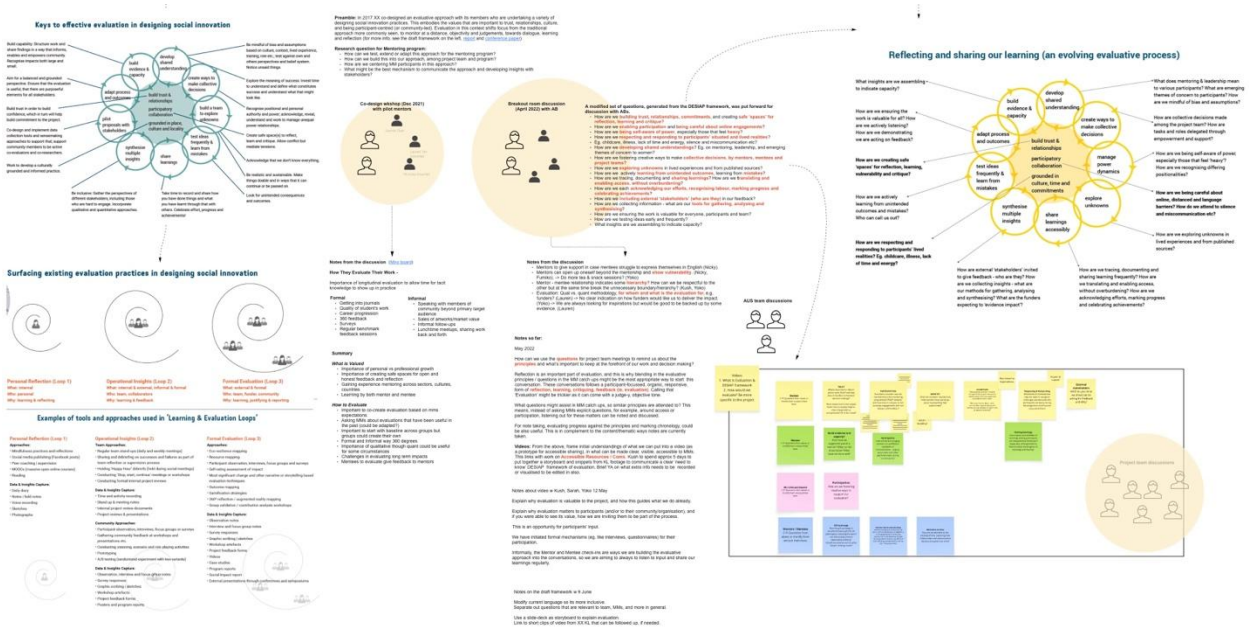


Figure 3. Team's Miroboard tracking the development of the yellow-flower framework

Meeting notes, recordings, workshop outputs, and participant feedback were revisited in relation to the Core Values and Principles, allowing patterns, tensions, and shifts in practice to surface. We assessed progress collectively and as a standing agenda item in team meetings and in dedicated evaluation sessions held every three to six months. These sessions were oriented towards the question: where are we being true to our values, and where are we falling short? Traditional Evaluation can separate reflexivity from evaluative practices, by prioritising

neutrality, distance and quantitative measurement of outcomes at the end. In contrast, our process was simultaneously evaluative and reflexive, like two sides of the same coin, to avoid the erasure of the researchers and participants who are central actors in responding and adapting to the contingency of change.

Core Values, Principles and Questions to reflect and learn

The questions generated in relation to the *Core Values* and *Principles* were purposefully critical and helped us to be accountable. We start with the *Core Values* (central yellow circle):

- *Build trust and relationships:* Positioning trust and relationships as central helps not to take this for granted, sharpened further by the question: *How are we creating safe 'spaces' for reflection, learning, vulnerability and critique?*

- *Participatory collaboration:* This aims to build ownership, delegate tasks, invite input to achieve collective aims together. As the program commenced mostly online due to COVID-19 restrictions, we adapted the questions to reflect these contexts: *How are we being careful about online, distance and language barriers? How do we attend to silence and miscommunication etc?*

- *Grounded in culture, time and commitments:* People are always situated differently. To heighten our awareness that women were taking time from their professional and personal responsibilities to be involved in the program, we asked: *How are we respecting and responding to lived realities? Eg. childcare, illness, lack of time and energy?*

Nine Principles (outer circles) are:

- *Develop shared understanding:* Taking time to be mindful of one's own and others' experiences, perspectives and expectations can help collectively understand what constitutes 'success': *What does mentoring and leadership mean to us collectively? What are emerging themes of concern? How are we mindful of bias and assumptions?*

- *Create ways to make collective decisions:* Attending to different roles, capabilities, capacities, expertise, experiences and team dynamics is valuable. With long-term and new team members joining, these questions helped to guide team activities: *How are collective decisions made? How are tasks and roles delegated through empowerment and support?*

- *Manage power dynamics:* This principle replaced *'pilot proposals with stakeholders'* because we had already run a pilot. The new focus, however, maintains the importance of revealing, acknowledging and understanding personal and positional power, and working to manage the harm that can come from disempowering hierarchies. *How are we being self-aware of power, especially those that feel 'heavy'? How are we recognising differing positionalities?*

- *Explore unknowns*: There is much we do not know, to be explicit about uncertainties: *How are we exploring unknowns in lived experiences and from published sources?*

- *Share learnings accessibly:* Recording and sharing discoveries that is attentive to differences in language and time constraints, considers the trans-cultural and -sectoral nature of the program. *How are we tracing, documenting and sharing learning frequently? How are we translating and enabling access, without overburdening? How are we acknowledging efforts, marking progress and celebrating achievements?*

- *Synthesise multiple insights:* Gather diverse perspectives from participants, the Advisory Board and others. *How are external 'stakeholders' invited to give feedback - who are they? How are we collecting insights - what are our methods for gathering, analysing and synthesising? What are the funders expecting to 'evidence impact'?*

- *Test ideas frequently and learn from mistakes:* Look for and attend to unintended consequences and outcomes with co-designing data collection tools and sensemaking approaches. It supports community members to be active co-evaluators and co-researchers. *How are we actively learning from unintended outcomes and mistakes? Who can call us out?*

- *Adapt process and outcomes:* Modify plans on the fly and abandon prior decisions in response to continuous feedback. *How are we ensuring the work is valuable for all? How are we actively listening? How are we demonstrating we are acting on feedback'?*

- *Build evidence and capacity:* How is work structured and findings shared in a way that informs, enables and empowers the team and participants. *What insights are we assembling to indicate capacity'?*

The team used the yellow-flower framework regularly as a 'critical friend', for example by including it as a standing agenda item in program meetings. It provided a scaffold for individual and group reflections that informed our ongoing practice, as well as for interactions with participants through program huddles, mentor and mentee check-ins and Advisory Board discussions. The term 'evaluation' can evoke a sense of being assessed or judged and lead to defensiveness, so we chose to use 'reflection and learning', towards fostering more open and constructive communication. This approach aligns with the principle *'Share learnings accessibly'*. We also translated the framework into Japanese and Thai in our program handbook. Additionally, recognising that the diagram could be opaque without contextualisation and explanation, we walked participants through it step-by-step. As the program drew to a close, the team selected key questions for deeper reflection on program impact and undertook interviews with participants. Sharing the Miroboard and using the questions in these different ways helped us to collaboratively reflect and learn with participants throughout.

The third adaptation: reflection and learning with participants

Given the scale of documentation collated, we present two stories of adapting and transforming-with the framework. These are not to evidence how 'successful' the framework 'worked', but instead, to surface frictions, challenge assumptions and trace the changes we made that

characterise reflexive *frameworking*. In sharing what we aimed for and decisions we abandoned, we illustrate the dynamic and continuous nature of evaluative practice.

Evaluating and adapting how we build trust and relationships

Eating together has been an DESIAP 'methodology', to situate and build lasting relationships (Akama and Yee 2024). The program began with the team sending drinks and snacks from our respective cultures, accompanied by personal stories about their significance, amongst the team and to participants. We invited them to enjoy the treats during online meetings, aiming to foster a sense of care and encourage light-hearted yet intimate conversations, helping us connect more deeply. This reflected a common custom of gift-giving when embarking on new relationships and hoping for an auspicious connection. This Gift Exchange became the first phase of relationship-building activity during the program.

We then invited the participants to exchange gifts within their own mentoring groups of three to four. We suggested something inexpensive and easy to post internationally, like the drinks and snacks we have set out, but refrained from specifying the items and cost. However, during early check-ins with the participants, we learned that the pressure and uncertainty of selecting gifts was causing stress for some (Akama et al 2024). This was a significant mistake we made, which had the potential to derail relationships we were trying so hard to foster early on. An activity that was designed as an expression of care, produced the opposite effect, causing discomfort and stress for some. Our intention to be participatory place an unexpected burden of cultural navigation on the participants.

Turning to the principle, *'Test ideas frequently and learn from mistakes'*, and the question, *'How are we actively learning from unintended outcomes and mistakes?*' helped us pause the activity and evaluate what we were missing. The participants inside the mentoring cluster were not well-acquainted when invited to join the Gifts Exchange, and most importantly, there were age, experience and seniority differences between them. Gifts exchanged between individuals of different social standing often require careful consideration to avoid perceived disrespect or imbalance (Mullis 2008). When examining gift-giving rooted in historical, social and philosophical traditions in different cultures, expected yet varying cultural norms guide the thought, value and symbolism that go into selecting a gift. These norms steer the roles, including how the receiver reciprocates. These interconnected social obligations are important to consider, since not doing so can place emotional burden on the receiver (Schrift 2014). These informed our observation of hesitation amongst younger mentees who may have wanted to avoid offending their senior mentors by prematurely giving an inappropriate gift. Reflecting on the principle, *'Managing power dynamics'* and the question *'How are we recognising differing positionalities?'* prompted us to be more mindful about how hierarchy impacts gift-giving. We assured participants that they could join the Gift Exchange as and when they felt ready to. Many did just that. Some groups waited until they felt they could give more meaningful, personal gifts. Some worked the Gift Exchange into mentoring activities as a form of making and collaborating, to learn and be inspired by each other's creativity. This enabled the Gift Exchange to become a longer participatory activity, and undoubtedly shifted how we built trust and relationships in the program. It also signaled early to the participants how we, as a team, were listening and open to modifying ideas, outcomes and schedules, based on their input.

Evaluating and adapting how we build 'capacity'

Building capacity is a desired and expected aim for social innovation initiatives (Alijani 2016; Pulford 2018). We too aimed to build capacities in the mentees, like leadership, based on individual strengths and opportunities. However, transcultural group mentoring was a new experience for all. This meant at the start, the team put considerable effort into matching needs and experience in curating the mentoring clusters. We prioritised what mentees had said they preferred or expected when they joined the program, which in turn, became our focus and key indicators of growth and capacity being built. The program also aimed for mentees to increase agency and confidence so they can mentor those in their workplace or community towards the end of the program. We saw this as a potential 'scaling' of social impact beyond the project timeframe. As the mentoring program gathered momentum the team continued to review team meeting notes, recordings of participant check-ins, and ephemeral records of engagement with the participants (e.g. WhatsApp chats) for indications of growth in capacity. This followed the principle, *'Synthesise multiple insights',* to respond to the question, '*What insights are we assembling to indicate capacity?'*. This helped us look beyond commonly-used benchmarks such as career advancement, job promotion or increased responsibilities at work as outcomes of capacity building, to instead look deeper to recognise internal shifts in beliefs and values as meaningful indicators of growth for the mentee.

During a check-in, Australian mentor L expressed concerns about her ability to support a Thai mentee. L attributed this to cultural and language differences. Two things then happened. Firstly, the Thai mentee invited their friend into the mentoring cluster. This was initially to help with Thai-English translation, however it became evident that her friend's presence helped the mentee feel at ease. The proactive initiative indicated the mentee's capacity to responsively problem-solve. Secondly, as their group dynamic coalesced, J, the other mentee, played an active

role in supporting her fellow mentee. Their mentor, L, recognised this: '*...there were many moments where you (J) were providing advice and guidance to them in a way that I just could not have. Because you just had a much more nuanced understanding of a very kind of particular cultural thing she was referencing about the way things work.'* This was a key learning moment for the team. It indicated that mentors and mentees alike harnessed their assets to support one another to mutually develop unexpected capacities.

The question, '*How are we ensuring the work is valuable for all?*' further broadened our understanding of capacity. Collaboratively evaluating our practices with mentors through regular check-ins allowed us to recognise the impact the program was having on mentors too. One mentor expressed how much they were learning from their cluster: '*I feel like I'm getting as much back, definitely as I'm giving*'. Mentor N, an arts developer and manager, described how her interactions with the mentees expanded her understanding of artists and their creative practice. She could draw upon this knowledge to enrich other areas of her work.

As similar patterns emerged across other mentoring clusters, we began to understand how capacity building was distributed and collective, rather than the focus we had initially on individual mentees' development. We also learnt that evaluating and adapting how we build capacity requires us to abandon expected trajectories and outcomes. Again, the focus on people beyond the mentees broadened as we listened, observed, reflected and learnt together. Evaluative questions like, '*How are we actively listening? How are we recognising differing positionalities?*', became a running script in the background to help us identify that those on the team, were also variously building confidence, gaining knowledge, and learning ways to apply, adapt and problem-solve. This helped us recognise latent capacity beyond the mentoring clusters,

that included the program team, guests and Advisory Board, who were all part of the assets for growth and transformation.

In exploring what 'capacity' meant, the framework questions prompted us to evaluate how we (the team and participants) understood and defined 'capacity', and to explore how capacity emerges differently for different individuals. Rather than serving individual needs, emphasising skill-building through training materials or providing formal guidance on leadership, we were able to adjust and attend to many kinds of supportive relationships and its potential to nurture trust and favourable conditions for growth and transformation. We also abandoned major plans and outcomes, such as expecting mentees to mentor others by the end of the program. Most mentees didn't feel ready, or compelled to do this, so we abandoned that plan too. Instead, we assured everyone that we supported and trusted all to learn and grow at their own pace.

From an external framework to a reflexive 'frameworking'

Across three cycles of development, the framework was used simultaneously as an evaluative tool and a reflexive scaffold. Throughout the process, we checked the program against our values and refined our methods. Evaluative judgements (e.g. what was working, what needed changing, and what needed to be abandoned), emerged through these relational, collaborative conversations and not at an arms distance. This departure from Traditional Evaluation is a deliberate epistemological stance. We hope we have shown that evaluation, reflection and reflexivity as inseparable and need to be so, when designing and participating in social change. Our evaluative practice was embedded in the relational work of the program precisely because

the things we were trying to evaluate (trust, capacity, cultural safety, power dynamics) cannot be assessed from the outside but surfaced through ongoing engagements.

We noted previously that building evaluative capacities integrated into co-designing practices requires time and adaptation. This has been the case for the team too: this paper spans six-years of collaboration. Specific questions in the framework helped us reflect more deeply, individually and collectively. Insights from this iterative process guided real time adaptation. The open-endedness of the questions reduced the risk of over-hastily resolving or flattening experiences, and gave us confidence that we could stay true to our *Core Values* when adjusting pace, direction and outcomes on the fly.

Here, another caution. Our account might imply dependency or mere affirmation of using the framework will result in a 'positive' outcome. Indeed, the framework featured in key aspects of program activity and gatherings, however, whether we consistently used it or not mattered little overall. The presence of the framework early on was enough to both elevate and continue reflection and learning, and like training wheels on a bike, it helped to build habitual practice through time and focus dedicated to this work. The same could be said about the *Core Values,* like *'Building Trust and Relationships*'. The explicitness of 'naming' what is important to the program helped people who were both new to and familiar with DESIAP and the program to know and notice situations that contrasts with and contradicts them. Yet, knowing the harm that can come from expecting women to speak out when they might not feel comfortable to, required affective and emotional labour. This became informal and frequent check-ins, often beyond working hours, which also came with guilt of time away from family (Yee et al 2024). This anecdote reinforces how any values, principles, aims, declared in any framework requires

ongoing effort, attention and adaptation to navigate and respond to the uncertainties of change and contingencies together.

A conundrum for Service Design: to be or not to be a framework…

We hope our stories resonates. Wary of design's universalising tendencies we will not summarise the 'usefulness' of the framework. Nor will we try to persuade others to 'apply' it in their own projects, or substitute for external evaluation. Rather than 'using' our framework, we recommend readers to nurture practices of reflexive *frameworking* by co-designing their own evaluative approaches with their team and participants. We have aimed to show that frameworks are only as good as how they are used, by whom, when and why (Irani, Dourish and Mazmanian 2010; Light and Akama 2012). As we have shared, people do the work of caring, listening, doubting, questioning, reflecting and learning. This work occurs as relationships grow and change. This is one key takeaway.

Building confidence and collaborative experience of developing one's own evaluative and reflexive practice may go some way to address a sense of disempowerment to the pre-determined measures enforced by commissioning authorities. While such efforts may not be as significant to replace Traditional Evaluation approaches preferred by funding agencies, the attempt to do so can open a different conversation. Indeed, some funding agencies welcomed the opportunity to explore alternative approaches with us, especially those that were equally concerned by the distance, disconnect and disempowerment caused by their normative evaluation methods (Akama, Yee, Hill and Tjahja 2019). However, even though they identified the limitation of their traditional approaches, many organisations were uncertain about how to shift

their evaluative processes to support communities to track social change. This, then, is a ripe opportunity for contribution and co-designing.

Frameworks will always oversimplify dynamics and conditions. Instead, by demonstrating ways to build evaluative practices in (service) design and research, we invite others to co-design, iterate and adapt ways of scaffolding learning and reflexivity into that work for themselves, their team and their participants. This could start by identifying *Core Values* and *Principles* that can welcome the conditions and infrastructures of collaboration. Generating critical questions that address the complexities of change together can also be valuable. Throughout this work, allowing for iteration or abandonment when required is crucial. Viewing them as 'living' accompaniments – *frameworking* – can be one way to remain true to the emergent uncertainties of social change.

Acknowledgements We acknowledge unceded Woi Wurrung and Boon Wurrung sovereignty where some co-authors live and work on Kulin lands now called Melbourne, and pay our respect to their Elders, past and current. Our respect extends to all Traditional Custodians of the lands, waters and skies where this paper is presented, read and reviewed. We express our heartfelt gratitude to all DESIAP participants for their trust and generosity. An intention to collaborate respectfully underpinned all the activities and ideas that fed the writing, but the paper follows academic authoring conventions in naming the core team, listed alphabetically. No generative AI tools were used in the writing or preparation of this article.

Disclosure Statement The authors report there are no competing interests to declare. Ethics approval was granted by Northumbria University, UK. Subjects provided informed written consent.

Funding This work was supported by Arts and Humanities Research Council (UK) and Australia-Japan Foundation.

References

Akama, Yoko, Juliette Anich, Areli Avendano Franco, Jaz Choi, Liam Fennessy, Marius Foley, Leah Heiss, Tania Ivanka, Emma Luke, Rowan Page, Tanja Rosenqvist, Juan Sanin, Linus Tan. 2022. "Why Are We Here? A Reflexive Story of Decolonising." In *Proceedings of the Participatory Design Conference 2022 - Volume 1*, 26–35. Newcastle upon Tyne, UK: Association for Computing Machinery. https://doi.org/10.1145/3536169.3537777.

Akama, Yoko. Ko-Le Chen, Hirotake Imanishi, Yuko Kikuchi, Sarah Kushinsky, Sarah Teasley, Kemmiga Teerapong, & Joyce Yee. 2022b. "สตรี (SATREE): Supporting women in art, crafts and social innovation" [Public Presentation]. Jim Thompson Art Center, Bangkok, Thailand.

Akama, Yoko. Ko-Le Chen, Hirotake Imanishi, Yuko Kikuchi, Sarah Kushinsky, Sarah Teasley, Kemmiga Teerapong, & Joyce Yee. 2024. "Recasting 'shadow' expanding respectful hierarchies in participatory design practices". *Proceedings of the Participatory Design Conference 2024 – Volume 1,* 52-62. Sibu, Malaysia: Association for Computing Machinery. https://doi.org/10.1145/3666094.3666096

Akama, Yoko, and Alison Prendiville. 2013. "Embodying, Enacting and Entangling Design: A Phenomenological View to Co-designing Services." *Swedish Design Research Journal* 1 (1): 29–41.

Akama, Yoko & Joyce Yee. 2024. "Introduction: Inter-related worldviews of designing social innovation", edited by Yoko Akama and Joyce Yee, in *Entanglements of designing social innovation in the Asia-Pacific*, Chapter 1, 1-20. London: Routledge.

Akama, Yoko, Joyce Yee, Regina Hill, & Cyril Tjahja. 2019. *Impact and Evaluation in Designing Social Innovation: Insights from the DESIAP KL Workshop & Symposium*. https://desiap.org/wp-content/uploads/2019/06/DESIAP-Impact-Evaluation-Report-2019.pdf

Alijani, Sharam, Alvaro Luna, Javier Castro-Spila, and Alfonso Unceta. 2016. "Building Capabilities through Social Innovation: Implications for the Economy and Society" In *Finance and Economy for Society: Integrating Sustainability*, edited by William Sun, 293–313. Bingley, United Kingdom: Emerald Group Publishing Limited.

Berrocal, Beatriz Bonilla, Annalinda De Rosa and Valentina Auricchio. 2025. "What Can Design Do for Social Impact Measurement? Bridging Evaluation and Creative Practice." In *ServDes 2025: Empowering Diversity, Nurturing Lasting Impact*, edited by Ravi Mahamuni and Prasad Onkar, 6–10 October, Hyderabad, India. https://doi.org/10.21606/servdes2025.71.

Bornmann, Lutz. "What Is Societal Impact of Research and How Can It Be Assessed? A Literature Survey." *Journal of the American Society for Information Science and Technology* 64 (2): 217–233. https://doi.org/10.1002/asi.22803.

Bossen, Claus, christian Dindler, and Ole Sejer Iversen. 2016. "Evaluation in Participatory Design: A Literature Survey" In *Proceedings of the 14th Participatory Design Conference: Short Papers,*

Interactive Exhibitions, Workshops - Volume 2 (PDC '16), 151–160. New York: Association for Computing Machinery. https://doi.org/10.1145/2940299.2940303.

Brandsen, Taco, Sandro Cattacin, Adalbert Evers, and Annette Zimmer. 2016. "The Good, the Bad and the Ugly in Social Innovation." In *Social Innovations in the Urban Context: Nonprofit and Civil Society Studies: An International Multidisciplinary Series*, edited by Taco Brandsen, Sandro Cattacin, Adalbert Evers, and Annette Zimmer, 303–311. Cham, Switzerland: Springer International Publishing.

Corsetti, Lea. 2022. "Decolonising Evaluation Practice: Is it time for an increased conversation in Europe?" *European Evaluation Society; Evidence and Evaluation Forum,* July 13, 2022. https://europeanevaluation.org/2022/07/13/decolonising-evaluation-practice-is-it-time-for-an-increased-conversation-in-europe/

Corsten, Niels, and Jules Pick. 2019. "Service Design Maturity Model". *Touchpoint,* Volume 10 (3). Accessed June 9, 2026 https://medium.com/touchpoint/the-service-design-maturity-model-84e0b8c82cec.

Cousins, Bradley, and Elizabeth Whitmore. 1998. "Framing Participatory Evaluation." In *Understanding and Practicing Participatory Evaluation*, edited by E. Whitmore, 3–25. New Directions for Evaluation, no. 80. San Francisco, CA: Jossey-Bass.

Cünha, Jorge, Wellington Alves, and Madalena Araujo. 2022. "Challenges of impact measurement in social innovation: Barriers and interventions to overcome 1" *Revista de administração Mackenzie* 23 (6):1-32. https://doi.org/10.1590/1678-6971/eRAMD220077.en.

Drain, Andrew, Aruna Shekar, and Nigel Grigg. 2021. "Insights, Solutions and Empowerment: a framework for evaluating participatory design". *CoDesign*, *17*(1), 1–21. https://doi.org/10.1080/15710882.2018.1540641

Evans, Rosie, Adele Leah, and Emina K. Petrović. 2020. "Spaces of Empowerment: Shaping Inclusive Public Places through Decolonising Participatory Design in Aotearoa New Zealand." In *Imaginable Futures: Design Thinking, and the Scientific Method. Proceedings of the 54th International Conference of the Architectural Science Association*, edited by Ali Ghaffarianhoseini, Amirhoseini Ghaffarianhoseini and Nicola Naismith, et al., 996–1005. Auckland, New Zealand: The Architectural Science Association (ANZAScA).

Foglieni, Francesca. 2025. "Measuring the Impact of Service Design." Filmed February 25, 2025. ServDes'25 Conference Lecture Series video, 01:26:20. https://www.youtube.com/watch?v=OBWgBkCeExA.

Foglieni, Francesca, Beatrice Villari, and Froukje S. Visser. 2018. "About Evaluation in Service Design: As It Is and How It Could Evolve." *Service Design Proof of Concept. Proceedings of the ServDes 2018 Conference* 150 (39): 489–499. Linköping: Linköping University Electronic Press. https://ep.liu.se/ecp/150/039/ecp18150039.pdf.

Groves, Leslie C., and Rachel B. Hinton. 2004. *Inclusive Aid: Changing Power and Relationships in International Development*. 1st ed. Sterling, UK: Earthscan Routledge.

Hui, Allison. 2023. “Situating Decolonial Strategies within Methodologies-in/as-Practices: A Critical Appraisal.” *The Sociological Review* 71 (5): 1075–94. https://doi.org/10.1177/00380261231153752

Middernacht, Zdena, Andres N. Lluch, Wouter de Iongh. 2023. “Decolonising Monitoring and Evaluation: From Control to Learning.” *INTRAC*. https://www.intrac.org/decolonising-monitoring-and-evaluation-from-control-to-learning/.

Krasnopolskaya, Irina, and Irina Korneeva. 2020. “Social Innovation Measurement: A Room for Quantitative Metrics.” *International Review of Applied Economics* 34 (5): 567–82. https://doi.org/10.1080/02692171.2020.1776686.

Kah, Sally., and Temidayo Akenroye. 2020. “Evaluation of Social Impact Measurement Tools and Techniques: A Systematic Review of the Literature.” *Social Enterprise Journal* 16 (4): 381–402. https://doi.org/10.1108/SEJ-05-2020-0027

Kaplan, Robert. S., and David P. Norton. 1992. “The Balanced Scorecard: Measures That Drive Performance.” *Harvard Business Review* 70 (1): 71–79.

Lamping, Sally, Julian Chen, Toni Dobinson, Paul Mercieca, Sonja Kuzich, and Stephanie Dryden. 2024. “’Noted, But Not Interested’: Critical Collaborative Autoethnography and the Local Politics of Belonging.” *CoDesign* 22(1): 1-14. https://doi.org/10.1080/15710882.2024.2445304.

Light, Ann, and Yoko Akama. 2012. “The Human Touch: Participatory Practice and the Role of Facilitation in Designing with Communities.” *Proceedings of the Participatory Design Conference: Research Papers* 12 (1): 61–70. https://doi.org/10.1145/2347635.2347645

Mullis, Eric. C., 2008. “Toward a Confucian Ethic of the Gift.” *Dao* 7: 175–94. https://doi.org/10.1007/s11712-008-9049-x. .

Nguyen, Quynh, Emma Jaspaert and Birgit Harthum. 2024. “Talk the Talk, Walk the Walk: A Case Study on Systematic Evaluation in Participatory Design." *Proceedings of the Participatory Design Conference 2024: Volume* 1: 133–144. https://doi.org/10.1145/3666094.3666103.https://doi.org/10.1145/3666094.3666103https://doi.org/10.1145/3666094.3666103

Patton, Michael Q. 2011. *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*. New York: Guilford Press.

Pulford, Louise. 2018. “Networks and Capacity Building in Social Innovation.” In *Atlas of Social Innovation*, 212–215. https://www.socialinnovationatlas.net/fileadmin/PDF/einzeln/04_Future-Challenges-and-Infrastructures/04_06_Networks-and-Capacity-Building-in-SI_Pulford.pdf

PWC. 2019. Global Consumer Insights Survey 2019. Accessed October 8, 2025. https://www.pwc.com/ee/et/publications/pub/Consumer-Insight-Survey-2019-report.pdf .

Reichheld, F. Fredrick, and Rob Markey. 2011. *The Ultimate Question 2.0: How Net Promoter Companies Thrive in a Customer-Driven World*. Boston, MA: Harvard Business Press.

Schrift, Alan D. 1997. *The Logic of the Gift: Toward an Ethic of Generosity*. New York: Routledge. https://doi.org/10.4324/9780203760369.

Shaw, David M. 2015. “Blinded by the Light: Anonymization Should Be Used in Peer Review to Prevent Bias, Not Protect Referees.” *EMBO Reports* 16 (8): 894–897. https://doi.org/10.15252/embr.201540943.

Sheppard, Benedict, Garen Kouyoumjian, Hugo Sarrazin, and Fabricio Dore. 2018. *The Business Value of Design*. McKinsey and Company. Accessed October 8, 2025. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-business-value-of-design#/.

Sweeting, Ben, and Michael Hohl. 2015. “Exploring Alternatives to the Traditional Conference Format: Introduction to the Special Issue on Composing Conferences.” *Constructivist Foundations* 11 (1): 1–7. http://constructivist.info/11/1/001

Tuhiwai Smith, Linda, 2021. *Decolonizing Methodologies: Research and Indigenous Peoples*. 3rd ed. London: Zed Books. https://doi.org/10.5040/9781350225282.

Mertens, D. M. 2009. *Transformative Research and Evaluation*. New York: The Guilford Press.

Yee, Joyce, Yoko Akama, Areli Avendano Franco, Ko-Le Chen, Hirotake Imanishi, Yuko Kikuchi, Sarah Kushinsky, Nikki Ralston, Sarah Teasley, Kemmiga Teerapong,.2024. “Invisible Designing: Emotional and Affective Labour in Relational Participatory Practices”. *Proceedings of Participatory Design Conference 2024, Volume 2*: 67-72. Sibu, Malaysia ACM. https://doi.org/10.1145/3661455.3669872

Yee, Joyce, Yoko Akama & Kemmiga Teerapong. 2020. “Being community and culturally-led: Tensions and pluralities in evaluating social innovation.” Edited by Yoko Akama, Liam Fennessy, Sara Harrington, & Anna Farago, *Proceedings of ServDes.2020 Tensions, Paradoxes, Plurality*, Melbourne Australia, 458–471: Linköping University Electronic Press.

Yee, Joyce, Kemmiga Teerapong & Yoko Akama. 2023. “Calling on ‘aunties’ and ‘nieces’ empowering women in creative sectors in Southeast Asia through designing mentorship.” Edited by Yoko Akama and Joyce Yee in *Entanglements of designing social innovation in the Asia- Pacific*, 94–114. London: Routledge. https://doi.org/10.4324/9781003244684-9

Weaver, Paul M., and René Kemp. 2017. “A Review of Evaluation Methods Relevant for Social Innovation: With Suggestions for Their Use and Development.” TRANSIT Working Paper #14, TRANSIT: EU SSH.2013.3.2-1 Grant Agreement No: 613169.

Westcott, Michael, Steve Sato, Deb Mrazek, Rob Wallace, Surya Vanka, Carole Bilson, and Dianne Hardin. 2014. “The DMI Design Value Scorecard: A New Design Measurement and Management Model.” *Design Management Review* 24 (4): 10–16. https://doi.org/10.1111/drev.10257.