

# Murrayarra: Technical Report

Supporting the Evaluation of  
Digital Tools for Aboriginal and  
Torres Strait Islander Languages



FIRST  
LANGUAGES  
AUSTRALIA

# Supporting the Evaluation of Technology for Aboriginal and Torres Strait Islander Languages

Prepared by Linda Blake (FLA), Jennyfer Lawrence  
Taylor (ANU), Julian Vido (ANU), with input from Cathy  
Bow (CDU)

Technical Report: 30 March 2023

**Author Details:** Linda Blake (First Languages Australia), Jennyfer Lawrence Taylor (Australian National University) and Julian Vido (Australian National University)

School of Computing, College of Engineering, Computing, and Cybernetics  
The Australian National University, Canberra ACT 2600 Australia, [www.anu.edu.au](http://www.anu.edu.au)  
TEQSA Provider ID: PRV12002 (Australian University), CRICOS Provider Code: 00120C

and

First Languages Australia  
20 Kemp Street, Wallsend NSW 2287 Australia  
[www.firstlanguages.org.au](http://www.firstlanguages.org.au)  
[contact@firstlanguages.org.au](mailto:contact@firstlanguages.org.au)

© 2023 Australian National University and First Languages Australia

**Dedication:**

In memory of our dear friend and colleague Cathy Bow.

**Note:**

This content in this report has not been peer reviewed or accepted for publication at this time.

**Acknowledgments:**

We wish to acknowledge and thank the following people for contributing to this report (alphabetical):

Toby Adams, Robert Amery, Carolyn Barker, Faith Baidson, Emma Browne, Ben Foley, Murray Garde, Dianna Hardy, Knut Olawsky, Barbara Martin, Gulwanyang Moran, Bruce Pascoe, Annalee Pope and Coco Yu.

This project was jointly funded by the ANU School of Computing and First Languages Australia.

# Contents

Contents.....	3
1.0 Introduction .....	5
2.0 Related Work .....	6
2.1 Indigenous Evaluation in the Literature .....	6
2.2 Evaluation Approaches to Indigenous Projects in Human-Computer Interaction .....	7
2.3 Evaluation in Computer Assisted Language Learning (CALL) and Other Language-Related Computing Disciplines..	8
2.4 Other Evaluation Examples from Australian Projects .....	8
3.0 Methods .....	9
3.1 Project Context and Methods .....	9
3.2 Limitations and Areas for Future Work.....	9
4.0 Findings .....	11
4.1 Current Evaluation Practices and Perspectives.....	11
4.1.1 Current evaluation practices and approaches .....	11
4.1.2 Insights into what evaluation has offered project teams .....	11
4.1.3 Ideas for future evaluation approaches and activities.....	12
4.2 Metrics and Measures for Evaluating Digital Tools.....	12
4.2.1 Language learning and use.....	12
4.2.2 Technology use and engagement .....	13
4.2.3 Availability and access to digital tools.....	13
4.2.4 Strengthening language speakers' connection to their languages .....	13
4.2.5 Respecting cultural protocols in designing and using technologies for Indigenous languages.....	14
4.2.6 Increasing visibility and promotion of Indigenous languages .....	14
4.2.7 Longevity and sustainability of digital tools .....	14
4.2 Barriers to Evaluation .....	15
4.3.1 Lack of funding, human resources, and time .....	15
4.3.2 Competing priorities and issues within project teams.....	15
4.3.3 Political challenges to evaluation from funding bodies .....	16
4.4 Key issues and considerations for evaluation .....	16
4.4.1 Need for Aboriginal and Torres Strait Islander ownership of digital tools and evaluation activities.....	16
4.4.2 Respecting intellectual property rights .....	16

4.4.3 Remunerating Aboriginal and Torres Strait Islander peoples for language activities .....	17
4.4.4 Value of sharing insights from evaluation .....	17
5.0 A Resource for Evaluation Planning .....	18
5.1 Lessons and Recommendations for the Evaluation of Aboriginal and Torres Strait Islander Language Tools.....	18
5.1.1 Is evaluation is the right use of community resources?.....	18
5.1.2 Developing an evaluation plan.....	18
5.1.3 Accounting for different agendas in evaluation.....	18
5.1.4 Incorporating evaluation checkpoints in activities across a project .....	19
5.1.5 Addressing timelines for evaluation and impact beyond the current project and team .....	19
5.1.6 Sharing findings from evaluation activities .....	19
5.1.7 Addressing funding barriers and power imbalances in evaluation .....	19
6.0 Reflective Questions for Evaluation .....	20
Bibliography .....	22

# 1.0 Introduction

Evaluation is an activity that can be undertaken in a digital language tool project. Evaluation through the lens of human-computer interaction “involves collecting and analysing data about users’ or potential users’ experiences”, (Sharp et al., 2019, p.496) in order to “improve the [...] design” (ibid.). While evaluation has the potential to improve the tools available to Aboriginal and Torres Strait Islander communities and organisations, the literature review in this project indicates that little has been published on the evaluation of tools for Aboriginal and Torres Strait Islander languages, Looking internationally, Carpenter et al. (2021) also highlight the “evaluation gap” (p.136) in digital tools for Indigenous languages, noting that “there is little in the way of systematic and rigorous evaluation on the results of their use” (ibid). The authors discuss issues such as a lack of reporting on outcomes such as “language use impacts” from prior projects, the narrow timeline of technology projects providing a limited evidence base, and the need for “community-grounded” and engaged evaluation in collaborative projects (Carpenter et al., 2021).

Aboriginal and Torres Strait Islander language communities are interested in a range of language tools, such as websites, databases and dictionary apps. This project sought to understand evaluation practices in the development of language tools, including identifying barriers and opportunities for new approaches to support evaluation by project teams.

Building on work by project collaborators Bow (Bow, 2017b, Bow, 2020) Taylor (Taylor, 2020, Taylor et al., 2016), and First Languages Australia (2016, 2017) the aims of this project were to:

1. Understand current evaluation practices in technology design projects for Aboriginal and Torres Strait Islander language technologies, including identifying opportunities and barriers to evaluation; and
2. Develop a novel evaluation resource for use by Aboriginal and Torres Strait Islander communities and organisations to support the technology design process, and assess the impact of their digital tools.

## 2.0 Related Work

### 2.1 Indigenous Evaluation in the Literature

Evaluation is increasingly recognised as important in projects involving Indigenous communities. Worldwide, there is a growing recognition that evaluation in the context of programs impacting Indigenous communities has a history of representing the settler state and proceeding from a perspective of deficit-framing and satisfying external agendas of whether money has been well-spent (LaFrance and Nichols, 2008, Jacob and Desautels, 2013, Duke et al., 2021).

Scholarly literature on Indigenous evaluation advocates for:

- the importance of building evaluation practices on the existing knowledge systems and cultural values of the community (LaFrance and Nichols, 2008, Gresku et al., 2022);
- centring and prioritising Indigenous perspectives in evaluation to enable Indigenous ownership of the process and metrics (Jacob and Desautels, 2013, Duke et al., 2021);
- enabling sufficient time for the evaluation process to assess learning outcomes (LaFrance and Nichols, 2008, Price et al., 2012);
- ensuring community participation, including the involvement of Elders (LaFrance and Nichols, 2008); and
- evaluation serving as a process for building relationships throughout the community (Gresku et al., 2022).

Research to address the lack of Indigenous perspectives in the evaluation of projects impacting Indigenous communities includes:

- developing an evaluation framework about whether technology should be employed in the context of Indigenous language learning (Galla, 2016);
- evaluating the extent to which participatory and culturally sensitive approaches were being adopted in the evaluation of Aboriginal programs (Jacob and Desautels, 2013);
- evaluating mainstream children's health programs from Indigenous perspectives and making adjustments for Indigenous contexts (Murtha et al., 2021);
- evaluating Indigenous perspectives in chronic disease interventions (Gresku et al., 2022);
- introducing Indigenous led evaluation as an essential component of culturally appropriate approaches to governance and to ensure equity in Aboriginal and Torres Strait Islander health research (Duke et al., 2021);
- developing an Indigenous led and designed evaluation model in the United States context (LaFrance and Nichols, 2008); and
- conducting literature reviews on evaluation in areas such as Indigenous health (Gresku et al., 2022).

In the Australian context, government agencies have adopted embedding Indigenous perspectives into evaluation processes. Examples include the Department of the Prime Minister and Cabinet's "Indigenous Advancement Strategy Evaluation Framework" (Department of the Prime Minister and Cabinet, 2018) which is "a guide for evaluation of programs and activities under the Indigenous Advancement Strategy (IAS)" (Department of the Prime Minister and Cabinet, 2018, p.2), and the Productivity Commission's "A Guide to Evaluation under the Indigenous Evaluation Strategy", which seeks to centre "Aboriginal and Torres Strait Islander people, perspectives, priorities and knowledges" (Australian Productivity Commission, 2020, p.3). However, they do not necessarily provide appropriate frameworks for communities and organisations to evaluate the language tools that they create.

Indigenous communities and organisations are reclaiming evaluation processes to achieve their own goals, in line with their own epistemologies and values (e.g., LaFrance and Nichols, 2008). Some examples include the work of EvalIndigenous, which promotes the use of Indigenous-led evaluation approaches and methods (Indigenous Peoples in Evaluation, n.d.), and the National Collaborative Centre for Aboriginal Health (based in Canada), which has published a guide entitled "Indigenous Approaches to Program Evaluation" (National Collaborative Centre for Aboriginal Health, 2013).

However, there are limited accessible resources that Aboriginal and Torres Strait Islander communities in Australia can use to evaluate their own projects, with a particular focus on tools to support language maintenance and revitalisation. In the next sections, we consider existing evaluation frameworks and methods from computer science.

## 2.2 Evaluation Approaches to Indigenous Projects in Human-Computer Interaction

Human-computer interaction (HCI) is a discipline of computer science that involves “the design, evaluation and implementation of interactive computing systems for human use and [...] the study of major phenomena surrounding them” (ACM SIGCHI Curriculum Development Group, 1992, p.5). Since “evaluation is integral to the design process” (Sharp et al., 2019, p.496), particularly evaluation of the “usability” and “user experience” (UX) aspects of a design (ibid.), HCI/UX may offer relevant evaluation methods and tools for communities and organisations. Relevant resources include usability principles and heuristics (Nielsen, 2012, Nielsen, 1994), user experience dimensions (Usability.gov, n.d.), and accessibility guidelines (W3C, 2018).

Prior projects in HCI research with Indigenous peoples internationally highlights the ways that evaluation methods are being used in different social and cultural settings and in different stages of the design process. Projects such as Rodil and colleagues’ work with the San people in Namibia (Rodil et al., 2020), involved using virtual reality (VR) for relationship building and early-stage design work, providing a VR game for people to play and stimulating conversations about the user experience (Rodil et al., 2020). Jensen et al. worked with the Herero people in Namibia to create 3D visualisations to represent Indigenous knowledges through drawing activities, and demonstrations of prototypes (Jensen et al., 2012). Muntean et al. used interviews and observations of museum visitors to evaluate a tangible tabletop museum exhibit about Musqueam culture (Muntean et al., 2017). In some cases, evaluation addressed the design process itself, such as measuring student engagement in the Aboriginal Territories in Cyberspace Skins educational digital skills workshop, for school students in Canada (Lameman et al., 2010).

A small number of HCI projects have involved designing technologies to specifically address language documentation and revitalisation. Birrbay and Dhanggati language teacher and research, Gulwanyang Moran, focused on gamification and language acquisition in a Masters research project and has made her project research report (Moran, 2021) and literature review (Moran, 2021a) available for community reference. Moran (2021) examines whether gamification really helps with language acquisition and retention of language, and the role of gamification for language revitalisation. Moran (2021) explores the role of technologies in language education, revitalisation and reclamation, highlighting the strengths and limitations of digital tools within a cultural context in language revitalisation and reclamation. Another key example is Plimmer et al.’s work with the Penan people in Malaysia to evaluate a “tangible user interface” (i.e., a smart object) for the preservation of their object sign language (Plimmer et al., 2015). Their evaluation activities included community meetings and design demonstrations with Elders, and a study to compare children’s language skills before and after using the tool (Plimmer et al., 2015). Further work with the Penan people by Zaman et al. (Zaman et al., 2015) to test a computer game called the “Oroo Adventure Game” involved children playing the game, and then completing a survey to evaluate its usability and to measure how well they had learnt the signs from the game (Zaman et al., 2015). This demonstrates that evaluation can involve both informal conversations and activities with community members, as well as formal surveys or tests to measure the effectiveness of digital tools for language learning. Other work in HCI shares guidelines for conducting usability studies outside lab-based settings, such as work with communities in the Congolese rainforest to design mobile applications (Vitos et al., 2017).

Several HCI projects have involved technology design and evaluation with Australian Aboriginal and Torres Strait Islander communities and organisations, with a focus on working with people out in communities and on Country. To co-design a digital platform for nation-building dialogue, Akama et al. (2016) showed Wiradjuri people a sketch of a possible system design using recordings they had made, as a way of having conversations with Elders about issues relevant to the digital platform (Akama et al., 2016). Taylor et al.’s work with the Wujal Wujal community used prototype demonstrations and a pilot study to evaluate how a soft toy design could encourage children’s Kuku Yalanji language use (Taylor et al., 2020). Soro et al. used prototype demonstrations and co-design workshops to build relationships and evaluate a Digital Community Noticeboard system on Groote Eylandt, developing a method called “cross-cultural dialogical probes” (Soro et al., 2016). Hardy et al.’s work to co-design a language revitalisation application with games with the Gugu Badhun people in Far North Queensland involved focus groups and fieldwork to



test early ideas and obtain feedback on prototypes (Hardy et al., 2016). With few HCI projects in this area, and a focus on early design work in some studies rather than finished products (e.g., Peters et al., 2018, Leong et al., 2020, Akama et al., 2016), there is an opportunity for further work to develop evaluation methods and practices for language projects with First Australians.

## 2.3 Evaluation in Computer Assisted Language Learning (CALL) and Other Language-Related Computing Disciplines

Other relevant computing disciplines including Computer Assisted Language Learning (CALL) and Mobile Assisted Language Learning (MALL) have developed their own evaluation frameworks and methods.

The range of evaluation frameworks and tools for technology designers and users include:

- Approaches to help people to make assessments about language learning technologies in relation to environmental factors, financial factors, socio-cultural factors, time constraints, and/or the community context (Baloh et al., 2015, Galla, 2016);
- frameworks focused on the design and usability of the language technology (Baloh et al., 2015);
- frameworks addressing the suitability of technologies for the pedagogical environment (e.g., in a classroom setting) (Rosell-Aguilar, 2017); and
- frameworks that provide a set of sequential steps that could be followed to develop and complete an evaluation (Koole and Ally, 2006, Vavoula and Sharples, 2009).

Papadakis (2021) conducted a systematic review of tools for the evaluation of educational apps for children (Papadakis, 2021). Papadakis argues that the Lee and Cherner's (2015) rubric is "the most comprehensive instrument of evaluation quality and depth" (Papadakis, 2021, p.41). It includes evaluation criteria grouped within three "domains" - "instruction", "design", and "engagement" (Lee and Cherner, 2015, p.37). However, it appears that the cultural context in which a technology is being deployed was often treated as sub-consideration in the tools surveyed.

Other computing work has focused on evaluation in the context of work with Indigenous peoples and languages. Ward (2018) highlights that the motivations of less commonly taught language communities and endangered languages communities tend to differ from those of more commonly taught languages, and it is therefore appropriate to take alternative approaches to evaluation in this context. Thus, CALL materials in an endangered language context can have positive social benefits, which may outweigh their short-term contributions to language learning (Ward, 2004). Galla's (2016) "Techancy Framework for Language Revitalization" provides a resource for communities to evaluate whether technology is "appropriate" given their language situation and contextual factors (Galla, 2016, p.1140). Bird (2020) advocates for perspectives of evaluation beyond "reductionist accounts of the performance of system components" (Bird, 2020, p.3510). Instead, Bird approaches evaluation as a "a device for engagement" that can advance local agendas and promote "culturally meaningful collaborative work" (Bird, 2020, p.3510-3511).

This resonates with the literature on Indigenous approaches to evaluation, which places emphasising the importance of relationality and cultural context, including the need to draw on Indigenous knowledge systems and cultural practices (National Collaborative Centre for Aboriginal Health, 2013, Chilisa, 2019). New evaluation approaches and resources are needed to assist communities and organisations in this regard.

## 2.4 Other Evaluation Examples from Australian Projects

The literature review conducted in this project suggests that there is only a small number of examples of evaluation in technology design projects for, or involving, Aboriginal and Torres Strait Islander languages. Some examples of published work and what was learnt from evaluation includes: (Green et al., 2011, Mamtora and Bow, 2017, Bow, 2019, Bow, 2017a, Foley et al., 2018, Keane et al., 2019, Szapiro et al., 2020, Bettinson and Bird, 2021)

## 3.0 Methods

### 3.1 Project Context and Methods

This project involved “semi-structured interviews” (Sharp et al., 2019) with nineteen people who have been involved in technology design projects for Aboriginal and Torres Strait Islander languages. The interviews were conducted online using a “yarning” approach (Bessarab and Ng'andu, 2010) where possible. Interviewees included Aboriginal and Torres Strait Islander community members and organisations, linguists and language revitalisation researchers, and ICT researchers and ICT professionals.

The project team included:

- Linda Blake, Wiradjuri mother, grandmother, language revivalist and secondary school teacher of Aboriginal Languages, Digital Technologies and Food Technologies (Blake)
- Julian Vido, Italian-Australian research officer with professional experience as a lawyer (Vido)
- Jennyfer Taylor, British-Australian university lecturer in human-computer interaction (Taylor)

Questions targeted the participants' views on evaluation, the role (if any) evaluation had played in their design process, the use and sustainability of their digital tools, and what evaluation activities they might use in future for ongoing projects.

The interview recordings were transcribed and a collaborative process of “thematic analysis” (Braun and Clarke, 2006) was used to identify themes repeated in the discussions. An initial set of codes was developed by one of the team members reading through the transcripts of their interviews, which a second team member then applied to transcripts from interviews they had conducted, with additional codes added to the codebook. A third team member then reorganised the analysis to ensure that content on the same themes was brought together, checked the coverage across themes, and included additional key ideas or concerns. Based on the interview findings, we read through the data to develop some insights towards an evaluation resource to help project teams evaluate their digital tools. This process provided insights to guide the publication of a resource to assist project teams to evaluate their digital tools.

The project received ethics approval from the ANU Human Research Ethics Committee (Protocol 2021/599) and participants received a gift voucher in recognition of their contributions.

### 3.2 Limitations and Areas for Future Work

There were several limitations to this project's approach, which could be addressed in future work.

First, the literature review may not have located all of the relevant literature that addresses evaluation in technology projects for Aboriginal and Torres Strait Islander languages. In particular, there is a lack of literature by Indigenous authors. Our account of past projects was not exhaustive but provided literature on evaluation that can be found online using search terms such as “Aboriginal language digital tool”, “Indigenous technology design”, “Aboriginal digital tool evaluation” and “Indigenous language technology evaluation”. This also might not locate publications that address issues relevant to evaluating a tool but are not described as evaluation activities or findings.

Secondly, a small sample of people involved in technology design projects for Aboriginal and Torres Strait Islander languages were interviewed. As such, there may be additional perspectives or insights in the broader community that are not reflected in our analysis. We have tried to give broad coverage to diverse voices and projects in the literature surveyed and interview participants recruited, however future work could use a larger sample.

The original intent of the project was to develop an evaluation “framework” similar to the literature (Rosell-Aguilar, 2017, Lee and Cherner, 2015). However, it became clear that technology design projects for Aboriginal and Torres Strait Islander languages are diverse in their aims, contexts and approaches. The researchers recognised that it would be difficult for one framework to include specific dimensions that are meaningful across these different situations. Instead, the project has developed a “resource” that identifies key lessons for evaluation based on our research, offers reflective questions for communities and organisations to use in planning evaluation for their own projects, and

compiles some useful further readings on evaluation. This builds on prior work by First Languages Australia such as *Warra: Building teams, building resources* (First Languages Australia, 2015b), *Angety Map: Digital Resource Report* (2016b) and *Junyirri: Framework for planning community language projects* (First Languages Australia, 2015a).

Initially, co-design workshops to further elaborate on the reflective questions included in this resource were proposed. However, this activity was not completed during the project period.

This project largely adopted a Western perspective on evaluation. Indigenous peoples have their own research paradigms. Linda Blake and First Languages Australia have worked to guide the project in the hope that it is respectful of the aspirations of Aboriginal and Torres Strait Islander language communities. Project teams applying the framework should determine what evaluation means to them, and how they wish to draw on Indigenous and/or Western traditions to evaluate their digital tool designs.

## 4.0 Findings

This section shares findings from the interview data that summarising the perspectives of participants, and analysis and recommendations. Section 4.1 focusses on describing the current state of practice, while Sections 4.2 – 4.4 present the key themes and patterns from our thematic analysis. Each of the interviewees works as part of a team to produce language tools and these are referred to as "project teams" throughout this section.

### 4.1 Current Evaluation Practices and Perspectives

#### 4.1.1 Current evaluation practices and approaches

Perspectives on what evaluation means, and the purpose of evaluation within the design process, differed across projects. While some project teams had undertaken evaluation of their digital tools, others had not, or may not have considered their activities as such. The emphasis on evaluation has increased in recent times through evolving conversations within specific research communities. Methods or activities that project teams reported to have used for collecting feedback and insights about their tool included:

- community demonstrations of design
- observing people carrying out tasks or activities using the design
- showing and discussing early sketches
- informal conversations to seek anecdotal feedback
- counting the number of downloads of an online resource
- counting the citations of a resource
- tracking engagement through community contact with the project team to ask questions
- project meetings within teams and with collaborators
- analysing user data within the tool or use logs
- monitoring use of the tool within particular groups or organisations
- conducting feedback surveys

This shows that evaluation has been conducted using both established methods such as surveys, and through more informal approaches such as engaging in conversations with community members as the project progresses. This project argues that both approaches are valid for obtaining useful feedback and insights to inform digital tool design. Evaluation methods should be chosen that are meaningful within the context of a specific project. The interviews highlight the range of unpublished evaluation activities that are taking place.

Challenges in balancing competing priorities when determining how to acquit project funding were noted. The interviewees suggested that there is often a need to prioritise digital tool creation to build up a critical mass of resources, or digitisation of materials over evaluation. One participant advocated for an upfront assessment of what will be most effective for a particular language community. That is not to say that participants did not recognise the value and importance of evaluation in their work, just that they felt compelled at times to prioritise more generative projects.

#### 4.1.2 Insights into what evaluation has offered project teams

Interviewees discussed what they had learnt from either conducting formal evaluation activities, or from anecdotal conversations with community members. Project teams were noted as able to:

- identify how easy or difficult a technology design was to use
- better target the design to particular audiences such as children
- see what aspects of the design users were interested in
- identify errors (e.g., with the interface or content)
- improve the visual design of the tool so that it has the right appearance for that community
- gather ideas for new content.

The above points demonstrate the ways in which evaluation practices can help project teams to improve a technology design to ensure that it works well for a language community.

#### 4.1.3 Ideas for future evaluation approaches and activities

The interviewees also discussed what kinds of evaluation approaches were possible or of interest, in some cases contingent on have access to additional resources. Approaches identified included:

- analysis of session times (how long someone spends using a technology<sup>1</sup>)
- using eye tracking data (e.g., where people look when they use the technology<sup>2</sup>)
- analysing online reviews
- tracking learning within a tool
- reviewing aspects of the system design against established principles
- reflecting on the authenticity and integrity of way language is being taught and used with the tool
- identifying how well the digital tool is “working” or not for the community
- using more rigorous or systematic evaluation in upcoming projects
- gauging whether the design ‘meets user needs’.

It was suggested that language learning might be evaluated by reaching identified learning stages, in the same way that children reach specified milestones in their growth and development.

To conduct evaluation, interviewees stated that they would need to develop evaluation activities with the people involved in technology use and put evaluation criteria in place. A participant noted that they would benefit from having access to a template of “guiding questions” to help evaluation. The call for further resources to guide evaluation has provided some inspiration for the reflective questions that are provided in this report.

## 4.2 Metrics and Measures for Evaluating Digital Tools

Interview participants presented evaluation “metrics” (First Languages Australia, 2015a) or “measures of success” (Taylor et al., 2016) for the evaluation of digital tools, or for more general projects that could inform evaluation criteria. Participants described metrics that they had used themselves and/or that they thought were important for the evaluation of future projects.

### 4.2.1 Language learning and use

Language learning and increasing language use were an aim of several projects, or identified as relevant evaluation criteria for language tools. Successful language learning was represented by:

- a growth in language use within different contexts such as work and home
- increased proficiency or fluency of individual speakers
- the ability to hold conversations in a language
- being able to translate in ways that are consistent with Elders of a language.

Some projects targeted helping people to improve specific grammatical skills, indicating the need for evaluation metrics that address proficiency in the specific skills being targeted. This raises questions around how to measure language learning in ways that account for prior knowledge and differences in individual learning journeys. Some participants also discussed the ability of the design process itself to enable language learning to take place between people, suggesting that evaluation could also extend to the design activities.

---

<sup>1</sup> Explanation provided by the research team

<sup>2</sup> Explanation provided by the research team

#### 4.2.2 Technology use and engagement

Ongoing technology use and active engagement with a digital tool were seen as key measures of success. However, while some participants were aware of whether their digital tool was available or being used, other participants had limited visibility of the use of their digital tool. Some participants indicated that continued use was an indicator in and of itself that technologies were working well, meaning that evaluation in these cases was not necessarily needed or prioritised. Non-use of a digital tool was also seen as indicative of an unsuccessful project. One participant stated the following:

*“When you see a community using it, [...] that's the test that it's worked, and you should be evaluating that right through the project. And listening to the community” (Bruce Pascoe)*

Specific approaches to evaluating technology use included:

- whether the target audience is using the digital tool
- whether the tool has other uses beyond the primary purpose
- observing technology in use in different contexts and activities
- assessing number of downloads as an indicator of use.

Some participants talked about evaluating usability, and pointed out that digital tools will not be adopted unless they are quick and easy to load and use.

Interviewees expressed that the development of social practices around the tool were also seen as key to increasing language and/or technology use, including:

- embedding the use of tools in “structured learning” environments
- increasing language use by developing local capability in ways that align with the community’s values
- developing a community of practice around the tool
- holding community meeting to demonstrate the design.

Therefore, evaluation therefore needs to extend beyond the characteristics of the digital tool to consider how it fits with a language community’s practices, activities, and policies.

#### 4.2.3 Availability and access to digital tools

Ensuring that digital tools and the language data they hold is available and accessible to language speakers and other stakeholders was considered a key priority for many of our interviewees. It is important to note that communities have different views on whether to make digital tools publicly available or not. This suggests that these decisions need to be made by those language communities themselves. An interviewee shared that some tools also have the potential to make learning opportunities more readily available to language speakers to practice at home and ease pressure on language resources in community such as Elders’ time. Some projects sought to make resources widely available, including providing digital materials across a variety of formats and platforms. Thus, evaluation metrics could be developed that track whether the project is achieving this aim.

Participants also identified in keeping digital tools available, such as ensuring that subscription fees are paid, and responding to changes in the policies of mobile app providers. Digital tools must integrate with community infrastructure. Local considerations for designing digital tools include internet infrastructure and coverage and the devices and platforms that language speakers or households actually own and use. The development costs associated with particular platforms were noted, and there was a recommendation that language communities look to tools that do not require the purchase of additional equipment or input devices.

#### 4.2.4 Strengthening language speakers’ connection to their languages

Strengthening connection to language is a measure of success that goes beyond language and technology specific metrics, particularly for those impacted by the Stolen Generations.

Digital tools were perceived as ways to:

- help re-engage people with their language and build interest
- stimulate language memory
- reduce people's negative feelings about not knowing their language by having it "out there"
- assist with identifying family connections to language
- facilitate connections between Elders and younger generations of language speakers.

These kinds of outcomes and associated measures of success are not reflected in some generic frameworks for evaluating language learning technologies identified in the literature review.

#### 4.2.5 Respecting cultural protocols in designing and using technologies for Indigenous languages

Our participants described the importance of ensuring that projects respect cultural protocols are respected with regard to how languages should be taught and learned. Respectful protocols, depending in the project, include obtaining permission from language speakers for how language is taught in particular settings, working with a Reference Group of Elders ensures that protocols are followed for a language, and ensuring that communities have ownership of how language materials are used. The project team's analysis is that these values guiding overall projects need to extend to evaluation activities.

#### 4.2.6 Increasing visibility and promotion of Indigenous languages

Some participants expressed the view that recognition and awareness can be measures of success from Indigenous perspectives, resonating with the Second National Indigenous Languages Survey (Marmion et al., 2014). For example, one participant recounted the experience of researching their language and finding an account that stated it was an "extinct language", as follows:

*"When I read that this website said the Kullilli language is an extinct language, I couldn't accept that. I accepted that it might be asleep, but I wouldn't accept that it was extinct. So, the aim was to awaken it and to bring the mob together to see what we could do, because we knew that if we were able to get some of the old materials that were recorded and share them with people, we knew that people would start to remember things" (Toby Adams, Kullilli Language).*

Other participants suggested that the promotion of language can be a goal of language tools, and that resources should help to increase "prestige" of a language.

Analysis: Thus, the evaluation metrics should not be limited to language proficiency or technology use, but need to align with broader community objectives. Evaluating these dimensions requires assessing the impact of a tool on a community of language speakers as a whole, rather than being limited to individual use.

#### 4.2.7 Longevity and sustainability of digital tools

"Longevity" (a long lifetime) was highlighted as a key issue for communities. The longevity and security of language content within a tool was a measure of success. The security of data was also identified as a challenge, and in one case, data security was seen as a higher priority than the interface design. Some participants discussed the need to separate interface from content or materials for long-term data reuse, and to have data and design formats that will specifically allow for reuse. One participant described longevity of data and backwards compatibility (new version of software being able to read files made with older versions) as key criteria for choosing a new platform for redeveloping an existing digital tool.

Investing time and resources into technology projects only to encounter issues with ongoing access to the digital tool was reported as an issue. Format obsolescence (e.g., CDs, DVDs, and dated app software) was recognised as a challenge. Interviewees' views on stable or secure format for language data and digital tools differed, with conversations noting that:

- books are more durable than CDs or apps if stored in the right conditions
- data can be safeguarded by digitising it from tapes
- digital tools are more durable than print resources if they can be updated and upgraded

- online content sustains access more easily than content in apps.

Once again, this shows the diversity of language projects and contexts, and the fact that the same evaluation metrics and measures applicable in one context may not be relevant in others.

A key challenge is communities' ability to maintain technologies in the long-term. Participants voiced concern about long term maintenance arrangements and associated expenses, including communities not having access to funding for maintenance or sustaining designs. The need to update and improve tools as language knowledge and aspirations evolve was recognised, along with the need for long-term plans for handover of tools and updates. As such, evaluation needs to consider "future proofing", ensuring that language data and content can be leveraged for future projects even if interfaces and devices change.

## 4.2 Barriers to Evaluation

### 4.3.1 Lack of funding, human resources, and time

While there was the sense that evaluation can benefit projects, resource limitations were perceived to hinder the ability to conduct evaluation, particularly as an activity following the deployment of the digital tool. Participants identified a lack of funding as a fundamental barrier to evaluation, particularly after the conclusion of a funding period.

Some project teams and organisations reported a lack of technical expertise and reliance on external developers, meaning that an allocation of funds to software development was prioritised over evaluation activities such as focus groups or user testing. Participants also suggested that available funding tends to:

- focus on development rather than evaluation or maintenance of tools, or
- prioritise devices or formats as outputs.

This highlights the systemic barriers to evaluation posed by current funding models.

Analysis: The systemic barriers to evaluation posed by current funding models need to be overcome if communities are to evaluate their tools.

A lack of human resources was also identified as a barrier to evaluation. Some organisations described having limited personnel available to work with ICT developers on projects, with "sophisticated" evaluation methods as surveys requiring staff time that might not be available. In some collaborative projects, participants described logistical challenges with organising evaluation around community members' availability. Within project teams, evaluation can be driven or constrained by the enthusiasm and persistence of those involved and the availability of individual people driving evaluation. One interviewee talked about enlisting the support of younger generations for advice on what directions technology projects should take. However, our interviewees also highlighted the tenacity of communities in conducting evaluations using the resources to be found amongst long-standing community relationships, and regular conversations between project partners. Some communities and organisations are also eager to partner with universities to access additional resources for projects but are unsure of how to build those relationships.

Limited time for evaluation was a challenge related to funding. There can be limited time available to conduct evaluation for short/small projects, including when university students are involved. One organisation described needing to reallocate staff away from other projects or programs to prioritise working on an app. The pace of innovation cycles in digital technologies requires regular cycles of evaluation. Resourcing for evaluation therefore needs to span an appropriate time-period, such that its impacts (both positive and negative) for language learning and other outcomes can be assessed.

### 4.3.2 Competing priorities and issues within project teams

Different disciplines or people working on a project (such as language speakers, language teachers, linguists, project managers, IT developers, and programmers) may have different priorities for evaluation. Some participants identified challenges with different stakeholders having different evaluation goals and metrics, communication issues affecting the suitability of the final design for the community, and misaligned expectations in terms of what the project would



deliver. One interview stated that there is a need to align the goals and priorities of academic researchers and language workers in community organisations.

Evaluation needs to reflect the diverse goals of communities and projects. There may be different perspectives within a community about what tools should be developed for their language, or differing priorities for their languages.

Challenges to evaluation within communities may include:

- people not wanting to participate in, or be the subject of evaluation
- difficulties in eliciting critical perspectives about the design, such as identifying aspects that are not working well
- geographical distance between language speakers and other members of projects teams.

Some of the issues interviewees had experienced included:

- investing time and energy into collaborating, with the expectation of having a resource at the end only for it to be in formats that were not useful for the intended users
- the imposition of designs on Aboriginal communities by external companies
- providing language content to external organisations who then expected them to pay fees to access their language materials.

#### 4.3.3 Political challenges to evaluation from funding bodies

There were concerns about loss of current or future funding if the evaluation criteria relating to language learning or technology use are not met. Analysis: This suggests a need to revisit the ways that digital tool projects are being funded. As language learning takes a long time, evaluation results should be decoupled from funding decisions, allowing communities and organisations to assess the outcomes of their projects without jeopardising future resourcing.

Funding bodies also set their own requirements and metrics for reporting. These may include project management metrics (e.g., personnel, procedures, outputs, budgets), language-related metrics such as amount of content within a tool or number of languages served, demographics or composition of the project team, and whether particular artefacts are delivered. It is not clear how well these metrics align with the interests of language communities, or whether these metrics dictates what evaluation activities communities undertake in satisfying the reporting obligations of funders.

### 4.4 Key issues and considerations for evaluation

#### 4.4.1 Need for Aboriginal and Torres Strait Islander ownership of digital tools and evaluation activities

Participants advocated for community led/owned/controlled/driven approaches to technology design processes. This project proposes that this should include evaluation. There was an assertion that communities should be involved in the design process even if they do not have technical skills. Aboriginal leadership in the design process is needed to produce tools that are fit for purpose. From the perspective of the participants, technical people making a tool need to be in the "back seat" and advise but not drive the design process. A factor reported as contributing to unsuccessful projects was dependence on unreliable external parties.

#### 4.4.2 Respecting intellectual property rights

A strong theme to emerge from interviews with participants was Indigenous ownership of language content and IP in their tools. One interviewee shared their experiences of being "taken advantage of" for sharing materials in the past needing to pay to access that knowledge again. They suggested developing protocols to ensure sharing happens in the right way in future. Some communities would like to develop a business model to allow others to engage with their knowledge in ways that respect data sovereignty and provide funding for updates. Thus, placing an emphasis on ownership highlights the need for evaluation metrics that consider how the language content embedded in language tools is retained by the community the long-term.

#### 4.4.3 Remunerating Aboriginal and Torres Strait Islander peoples for language activities

There is a need to rethink approaches to community engagement in evaluation, particularly in collaborative projects between Aboriginal and Torres Strait Islander communities/organisations and external parties. Our participants made the point that community members must be appropriately remunerated for their expertise and input into language tool design activities. From the perspective of an interviewee, paying people for the work that they have done and covering travel costs is a way of showing respect to Elders. An interviewee also asserted that people “are the greatest tool”, with remuneration benefiting individuals and their communities.

Analysis: Why then should Aboriginal and Torres Strait Islander peoples be giving away their time and knowledge (as volunteers)? Indigenous people have community, family and public responsibilities in addition those experienced by other Australians, and participation in language projects can be a cultural responsibility. However, this responsibility does not reduce the need to work full-time to cover the cost of living and meet the plethora of other family and cultural commitments. Indigenous cultural projects contribute to the fabric of life for all Australians in ways that are both obvious yet also invisible to non-Indigenous persons. Financial reimbursement for time spent "volunteering" on language projects is the minimum that should be offered.

#### 4.4.4 Value of sharing insights from evaluation

Participants had mixed feelings about how insights from evaluation should be shared with other communities and project teams. On the one hand, some participants indicated a willingness to share evaluation findings and lessons with other project teams unless there are privacy and ethical concerns, pointing out the value in sharing their activities and lessons learnt. Evaluation insights were reported as being shared through academic publications and conferences such as PULiiMA: Indigenous Languages and Technology Conference and the Australian Languages Workshop.

On the other hand, a participant felt uncertain about communities’ desires to publish evaluation results, and noted that internally facing evaluation results might be more appropriate for their projects, or fed back to the developer to ensure the right approach is being taken. It was noted that a participant did not appreciate external parties evaluating their tools and publishing the results of external evaluation without including the perspectives of the people who built them.

## 5.0 A Resource for Evaluation Planning

### 5.1 Lessons and Recommendations for the Evaluation of Aboriginal and Torres Strait Islander Language Tools

#### 5.1.1 Is evaluation is the right use of community resources?

Project teams should consider whether undertaking evaluation is the right use of the available resources, and whether it can offer meaningful results. This challenges that idea that all projects should include evaluation. However, the literature and interviews provide examples of what evaluation can offer.

Evaluation must be led by Aboriginal and Torres Strait Islander people, or co-designed in partnership with them, so that they can choose whether and how to conduct evaluation and share results. Each Indigenous community has its own collective experience and choice of evaluation process, be it formal or informal, public or internal. This should not be restricted or enforced by funding organisations or other external collaborators. Respect is non-negotiable.

This position aligns with published principles for conducting and evaluating projects with Aboriginal and Torres Strait Islander people, such as “engagement with Aboriginal and Torres Strait Islander people” and “the importance of Aboriginal and Torres Strait Islander leadership” (Kelaher et al., 2018), “Indigenous self-determination” and “Indigenous leadership” (AIATSIS, 2020), and “centring Aboriginal and Torres Strait Islander people, perspectives, priorities, and knowledges” (Australian Productivity Commission, 2020).

#### 5.1.2 Developing an evaluation plan

Project teams might benefit from developing an evaluation plan to ensure the community needs and concerns are addressed throughout the design process. This plan might include formal and/or anecdotal evaluation methods.

Two resources that outline some elements of research or evaluation plans are:

- The *PRET A Reporter* framework for user studies in technology design projects (Blandford et al., 2008) which consists of “Purpose of evaluation”, “Resources and Constraints”, “Ethics”, “Techniques for gathering data”, “Analysis techniques”, and “Reporting of findings” (Blandford et al., 2008, p.9).
- The Queensland Government’s definition of “*evaluation plan*” (Queensland Government, 2020) for program evaluations, which could be adapted for digital tool projects. It describes evaluation plans as covering “the program’s objective, context and logic as well as the evaluation’s scope, objectives, principles, guiding questions, design, methodology, data collection, analysis and reporting” (Queensland Government, 2020, p.12).

Not all Aboriginal and Torres Strait Islander communities value non-Indigenous evaluation approaches. However, all Aboriginal and Torres Strait Islander communities are well placed to determine how to conduct evaluation in culturally safe ways that fit within their own context.

#### 5.1.3 Accounting for different agendas in evaluation

Digital tool projects are often undertaken collaboratively involving various individuals and organisations working together, as seen indicated in the interviews and Warra (First Languages Australia, 2015b). Different individuals within a project team may have their own perspectives on what should be evaluated and how. Project teams could consider addressing evaluation in discussions at the beginning of each project, to ensure evaluation aims and approaches are explicitly stated and reflect community priorities.

“Value” means different things to different stakeholders, and as such, a project’s success can only be evaluated against the aims or needs of that particular project. The values of the participants and target audience need to be implicit

within evaluation processes is necessary if quality data is to be collected and trusting relationships maintained. Aboriginal and Torres Strait Islander language speakers and owners are best placed to:

- determine whether Indigenous and/or non-Indigenous evaluation methods can offer them value
- identify culturally sensitive ways to solicit input into design processes
- collect data about technology use.

#### 5.1.4 Incorporating evaluation checkpoints in activities across a project

Evaluation does not need to be conducted as a separate and resource-intensive phase on its own at the end of a project. Evaluation can be *formative*, which is “conducted during design to check that a product continues to meet users’ needs” (Sharp et al., 2019, p.499) and/or *summative*, which is “carried out to assess the success of a finished product” (Sharp et al., 2019, p.500). If project teams do not have resources to carry out summative evaluation at the end of a tool project, they may consider whether formative evaluation questions or “checkpoints” can be built into the project design, within existing resources.

Ongoing evaluation with language reference groups or committees is championed as a means to continually gauge the extent to which the priorities of the community are being met (First Languages Australia, 2015b, First Languages Australia, 2015a). Building ongoing evaluation into a design process can help to ensure that there is strong community leadership and engagement. Evaluation might allow audiences to try things out and check that they are working as the process proceeds rather than as a burden at a project’s conclusion.

#### 5.1.5 Addressing timelines for evaluation and impact beyond the current project and team

The ongoing design, maintenance, and use of a tool may extend beyond the project timeframe and the commitment of the project team. If it would be useful to conduct evaluation activities after the tools have been released (e.g., six months or years into the future), then the project team could consider how to hand over responsibility for evaluation and evaluation planning materials. It is useful to consider how long a tool might work, and how speakers can make use of the language materials it contains once it is no longer available. Methods from HCI such as “multi-lifespan timelining” (Yoo et al., 2016) could help project teams to situate the current project with respect to past and future activity. While some evaluation methods take a Western linear view of time, Aboriginal and Torres Strait Islander projects may be able to adapt them to reflect their own perspectives.

#### 5.1.6 Sharing findings from evaluation activities

As few Indigenous digital language tool projects are publishing evaluations, there may be value in increased sharing of information and knowledge between communities and organisations. Some interviewees indicated a keenness to learn from and share with others. Sharing could take the form of written reports, conference presentations, videos, blog posts, or academic research articles. Some conference presentations can involve giving a talk to those present, but not making a recording or written record available. However, the reporting of evaluation activities should be a decision that communities and organisations make for themselves, and not enforced by funders or pressured by others.

#### 5.1.7 Addressing funding barriers and power imbalances in evaluation

Interviewees highlighted funding as linked to the other barriers to evaluation. Funding is linked to political agendas that determine how easy it is Indigenous people to access resources to implement and evaluate digital tool projects that are important to them. Funding was also seen as an indicator of respect as it enables the people involved in projects to be paid for their time and efforts in participating in community work and sharing their cultural knowledge. Communities should be able to access funding that supports the evaluation of the work they are undertaking, and the relevance of existing tools to their language contexts.

Political agendas can also misalign with community agendas. First Languages Australia reports (in conversations 2022) that Australia’s history has resulted in vacuum of trust between settler governments and Indigenous custodians. Further, Aboriginal and Torres Strait Islander community members can be hesitant to openly discuss issues or ask for help, concerned that they may open themselves or their community up to further government invasion or control. The

significance of trust is acknowledged by the Australian Government (Commonwealth of Australia, 2020). Decolonising approaches to technology design (Irani et al., 2010, Dourish et al., 2020, Bird, 2020) and research (Smith, 2012, Chilisa, 2019) recognise these power imbalances as present within the structures in which projects operate, and seek to address and disrupt uneven power relations. Power and trust need be accounted for when conducting evaluation.

## 6.0 Reflective Questions for Evaluation

Following are reflective questions that might help project teams to evaluate their language tools. These questions are based on the interviews and the project team's perspective. Questions consistent with concerns from academic literature are noted in the footnotes.

### Language learning and use

- Are individual learners' language skills improving as a result of using the tool?
- How well does the tool support language learners with differing skills and experiences?
- Are language speakers feeling more confident in learning and speaking the language by using the tool?
- Is language use by individuals, within families, and within the community increasing as a result of this tool?
- Is the tool supporting language learning in the way intended?
- Is the tool encouraging language use in ways that were not intended but are equally important as the intended outcomes?
- Is language use by individuals, within families, and within the community increasing as a result of this tool?

### Use of the tool

- Who is using your tool?
- How many people are using your tool?
- Are the intended people able to access the intended information? (e.g., is it child friendly<sup>3</sup>?)
- How often are people using your tool? How long are they spending when they use your tool? Do they use it more than once? Is the use of your tool increasing over time?
- Are there barriers preventing people from using the tool?
- How might users describe using the tool? (e.g., is it fun, annoying, clunky, easy?)
- What do users like or dislike about the design?
- Has the level of use of the tool made the effort/cost in its development worthwhile?
- Are users able to interact with the tool in ways that help them to learn?<sup>4</sup>

### Fitness for purpose and with existing infrastructure

- Does the tool fit with the environments, contexts, and infrastructures in which is intended to be used?<sup>5</sup>

### Increasing visibility and promotion of languages

- Does the tool or application make the language more visible?
- Does the tool assist with promoting or raising awareness of the language?<sup>6</sup>

---

<sup>3</sup> Reflects a similar concern to (Papadakis et al. 2020).

<sup>4</sup> See "Interactivity" in (Lee & Cherner 2015), (Rosell-Aguilar 2017).

<sup>5</sup> See (Galla, 2016) for the Technacy Framework.

<sup>6</sup> Resonates with (Ward & Genabith 2003)

- Have you used available channels to promote your tool (e.g., project website, social media, blog posts, wider media coverage)?

### **Longevity and sustainability of digital tools<sup>7</sup>**

- Has making the tool taken as long as expected?
- Is it still working?
- Can the language content be used again once the tool stops working?
- Does maintenance of the tool:
  - Require the same collaborators to continue working together? If not, how will information about the project be handed over to others?
  - Require specialist technical assistance (programming skills). If so:
    - Are there opportunities to train community members to develop the skills to maintain or evolve the tool?
    - Is there a plan for how maintenance and support will be paid for over the life of the tool?
- Can the language content be exported once the tool is no longer working?
- Thinking about how long the tool might last, has the investment been worthwhile?

### **Overall**

- What was achieved by developing the digital tool? Was producing this tool the best outcome for the community?
- Are there components of the tool (software, hardware, content) that can be reused by other communities or projects?<sup>8</sup>
- How might the learnings from this activity inform what tools are developed next?
- When could you evaluate your tool again?

---

<sup>7</sup> See (Baloh et al. 2015) for related concerns

<sup>8</sup> See (Ward and Genabith, 2003, Ward, 2004).

# Bibliography

- ACM SIGCHI CURRICULUM DEVELOPMENT GROUP 1992. ACM SIGCHI Curricula for Human-Computer Interaction. New York.
- AIATSIS 2020. AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research. Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS).
- AKAMA, Y., KEEN, S. & WEST, P. 2016. Speculative Design and Heterogeneity in Indigenous Nation Building. *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. Association for Computing Machinery.
- AUSTRALIAN PRODUCTIVITY COMMISSION 2020. A Guide to Evaluation under the Indigenous Evaluation Strategy. Canberra: Australian Productivity Commission.
- BALOH, M., ZUPANC, K., KOSIR, D., BOSNIC, Z. & SCEPANOVIC, S. A quality evaluation framework for mobile learning applications. 2015 4th Mediterranean Conference on Embedded Computing (MECO), 2015/06 2015. Budva, Montenegro: IEEE, 280-283.
- BESSARAB, D. & NG'ANDU, B. 2010. Yarning About Yarning as a Legitimate Method in Indigenous Research. *International Journal of Critical Indigenous Studies*, 3, 37-50.
- BETTINSON, M. & BIRD, S. 2021. Collaborative Fieldwork with Custom Mobile Apps. *Language Documentation and Conservation*, 15, 411-432.
- BIRD, S. 2020. Decolonising speech and language technology. *Proceedings of the 28th International Conference on Computational Linguistics*.
- BLANDFORD, A., ADAMS, A., ATTFIELD, S., BUCHANAN, G., GOW, J., MAKRI, S., RIMMER, J. & WARWICK, C. 2008. The PRET A Rapporteur framework: Evaluating digital libraries from the perspective of information work. *Information Processing & Management*, 44, 4-21.
- BOW, C. 2017a. Activating community-based Indigenous language and culture resources for university teaching: report on the development of a digital shell and pilot delivery. Darwin: Charles Darwin University.
- BOW, C. 2017b. Evaluating digital tools for endangered languages. *2017 Puliima Indigenous Languages and Technology Conference*. Cairns, Australia.
- BOW, C. 2019. Collaboratively designing an online course to teach an Australian Indigenous language at university. *Babel*, 54, 54-60.
- BOW, C. 2020. *Entanglements of digital technologies and Indigenous language work in the Northern Territory*. Charles Darwin University (Australia).
- BRAUN, V. & CLARKE, V. 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, 3, 77-101.
- CARPENTER, J., GUERIN, A., KACZMAREK, M., LAWSON, G., LAWSON, K., NATHAN, L. P. & TURIN, M. 2021. Locally Contingent and Community-Dependent Tools and Technologies for Indigenous Language Mobilization. In: LINK, A., SHELTON, A. & SPERO, P. (eds.) *Indigenous Languages and the Promise of Archives*. Nebraska: University of Nebraska Press.
- CHILISA, B. 2019. *Indigenous Research Methodologies*, Thousand Oaks, California, Sage Publications.
- COMMONWEALTH OF AUSTRALIA 2020. Indigenous Evaluation Strategy Productivity Commission Background Paper. Canberra: Australian Productivity Commission.
- DEPARTMENT OF THE PRIME MINISTER AND CABINET 2018. Indigenous Advancement Strategy: Evaluation framework. Canberra: Commonwealth of Australia.
- DOURISH, P., LAWRENCE, C., LEONG, T. W. & WADLEY, G. 2020. On Being Iterated: The Affective Demands of Design Participation. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- DUKE, D. L. M., PRICTOR, M., EKINCI, E., HACHEM, M. & BURCHILL, L. J. 2021. Culturally Adaptive Governance—Building a New Framework for Equity in Aboriginal and Torres Strait Islander Health Research: Theoretical Basis, Ethics, Attributes and Evaluation. *International Journal of Environmental Research and Public Health*, 18, 7943.

- FIRST LANGUAGES AUSTRALIA 2015a. Junyirri: A Framework for Planning Community Language Projects. First Languages Australia.
- FIRST LANGUAGES AUSTRALIA 2015b. Warra: Building Teams, Building Communities. First Languages Australia.
- FOLEY, B., ARNOLD, J. T., COTO-SOLANO, R., DURANTIN, G., ELLISON, T. M., ESCH, D. V. & AL., S. H. E. 2018. Building Speech Recognition Systems for Language Documentation: The CoEDL Endangered Language Pipeline and Inference System (ELPIS). 6th Intl. Workshop on Spoken Language Technologies for Under-Resourced Languages: SLTU.
- GALLA, C. K. 2016. Indigenous language revitalization, promotion, and education: function of digital technology. *Computer Assisted Language Learning*, 29, 1137-1151.
- GREEN, J., WOODS, G. & FOLEY, B. 2011. Looking at language: Appropriate design for sign language resources in remote Australian Indigenous communities. *Sustainable data from digital research: Humanities perspectives on digital scholarship*. University of Melbourne: Custom Book Centre.
- GRESKU, D., JONES, C. & KURTZ, D. L. M. 2022. Collaborative Evaluation Frameworks for Indigenous-Led Community Health Interventions: A Scoping Review. *Canadian Journal of Program Evaluation*.
- HARDY, D., FOREST, E., MCINTOSH, Z., MYERS, T. & GERTZ, J. 2016. Moving beyond Just Tell Me What to Code : Inducting Tertiary ICT Students into Research Methods with Aboriginal Participants via Games Design. *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. Association for Computing Machinery.
- INDIGENOUS PEOPLES IN EVALUATION. n.d.,. *EvalIndigenous* [Online]. Available: <https://comm.eval.org/aeaipetig/evalindig> [Accessed 19 Dec 2022].
- IRANI, L., VERTESI, J., DOURISH, P., PHILIP, K. & GRINTER, R. E. 2010. Postcolonial Computing: A Lens on Design and Development. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- JACOB, S. & DESAUTELS, G. 2013. Evaluation of Aboriginal Programs: What Place is Given to Participation and Cultural Sensitivity? *International Indigenous Policy Journal*, 4.2.
- JENSEN, K. L., WINSCHIERS-THEOPHILUS, H., RODIL, K., WINSCHIERS-GOAGOSES, N., KAPUIRE, G. K. & KAMUKUENJANDJE, R. 2012. Putting It in Perspective: Designing a 3D Visualization to Contextualize Indigenous Knowledge in Rural Namibia. *Proceedings of the Designing Interactive Systems Conference*. Association for Computing Machinery.
- KEANE, T., CHALMERS, C., BODEN, M. & WILLIAMS, M. 2019. Humanoid robots: learning a programming language to learn a traditional language. *Technology, Pedagogy and Education*, 28, 533-546.
- KELAHER, M., LUKE, J., FERDINAND, A., CHAMRAVI, D., EWEN, S. & PARADIES, Y. 2018. An Evaluation Framework to Improve Aboriginal and Torres Strait Islander Health. Melbourne: The Lowitja Institute.
- KOOLE, M. & ALLY, M. Framework for the Rational Analysis of Mobile Education (FRAME) Model: Revising the ABCs of Educational Practices. International Conference on Networking, International Conference on Systems and International Conference on Mobile Communications and Learning Technologies (ICNICONSMCL'06), 23-29 April 2006 2006. 216-216.
- LAFRANCE, J. & NICHOLS, R. 2008. Reframing evaluation: Defining an Indigenous evaluation framework. *The Canadian Journal of Program Evaluation*, 23, 13-31.
- LAMEMAN, B. A., LEWIS, J. E. & FRAGNITO, S. 2010. Skins 1.0: A Curriculum for Designing Games with First Nations Youth. *Proceedings of the International Academic Conference on the Future of Game Design and Technology*. Association for Computing Machinery.
- LEE, C.-Y. & CHERNER, T. S. 2015. A comprehensive evaluation rubric for assessing instructional apps. *Journal of Information Technology Education: Research*, 14, 21-53.
- LEONG, T. W., LAWRENCE, C. & WADLEY, G. 2020. Designing for Diversity in Aboriginal Australia: Insights from a National Technology Project. *Proceedings of the 31st Australian Conference on Human-Computer-Interaction*. Association for Computing Machinery.



- MAMTORA, J. & BOW, C. 2017. Towards a Unique Archive of Aboriginal Languages: A Collaborative Project. *Journal of the Australian Library and Information Association*, 66, 28-41.
- MARMION, D., OBATA, K. & TROY, J. 2014. Community, identity, wellbeing: the report of the Second National Indigenous Languages Survey. Canberra: Australian Institute of Aboriginal and Torres Strait Islander Studies.
- MORAN, G. 2021. Gathang Yawutung: Gamification of language learning for acquisition and retention of a sleeping language.  
<https://yaale.com.au/resource/Research%20Report:%20Gamification%20of%20language%20learning%20for%20acquisition%20and%20retention%20of%20a%20sleeping%20language>
- MORAN, G. 2021a. Discussion paper: Technology and Language Learning Assessment.  
<https://yaale.com.au/resource/Discussion%20paper:%20Technology%20and%20Language%20Learning%20Assessment>
- MUNTEAN, R., ANTLE, A. N., MATKIN, B., HENNESSY, K., ROWLEY, S. & WILSON, J. 2017. Designing Cultural Values into Interaction. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- MURTHA, K., THOMPSON, K., CLELAND, P. & GALLEGOS, D. 2021. Adaptation and evaluation of a nutrition and physical activity program for early childhood education settings in Aboriginal and Torres Strait Islander communities in remote Far North Queensland. *Health Promotion Journal of Australia*, 32, 163-171.
- NATIONAL COLLABORATIVE CENTRE FOR ABORIGINAL HEALTH. 2013. *Indigenous Approaches to Program Evaluation* [Online]. Available: <https://www.nccih.ca/docs/context/FS-IndigenousApproachesProgramEvaluation-EN.pdf> [Accessed 10 March 2023].
- NIELSEN, J. 1994. 10 Usability Heuristics for User Interface Design. Available from: <https://www.nngroup.com/articles/ten-usability-heuristics/> [Accessed 29 January 1994].
- NIELSEN, J. 2012. Usability 101: Introduction to Usability. Available from: <https://www.nngroup.com/articles/usability-101-introduction-to-usability/> [Accessed 29 January 2023].
- PAPADAKIS, S. 2021. Tools for evaluating educational apps for young children: a systematic review of the literature. *Interactive Technology and Smart Education*, 18, 18-49.
- PETERS, D., HANSEN, S., MCMULLAN, J., ARDLER, T., MOONEY, J. & CALVO, R. A. 2018. Participation is Not Enough : Towards Indigenous-Led Co-Design. *Proceedings of the 30th Australian Conference on Computer-Human Interaction*. Association for Computing Machinery.
- PLIMMER, B., HE, L., ZAMAN, T., KARUNANAYAKA, K., YEO, A. W., JENGAN, G., BLAGOJEVIC, R. & DO, E. Y.-L. 2015. New Interaction Tools for Preserving an Old Language. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- PRICE, M., MCCOY, B. & MAFI, S. 2012. Progressing the Dialogue about a Framework for Aboriginal Evaluations: Sharing Methods and Key Learnings. *Evaluation Journal of Australasia*, 12, 32-37.
- QUEENSLAND GOVERNMENT 2020. Program Evaluation Guidelines. Brisbane: The State of Queensland (Queensland Treasury).
- RODIL, K., MAASZ, D. & WINSCHIERS-THEOPHILUS, H. 2020. Moving Virtual Reality out of Its Comfort Zone and Into the African Kalahari Desert Field: Experiences From Technological Co-Exploration With an Indigenous San Community in Namibia. *26th ACM Symposium on Virtual Reality Software and Technology*. Association for Computing Machinery.
- ROSELL-AGUILAR, F. 2017. State of the App: A Taxonomy and Framework for Evaluating Language Learning Mobile Applications. *CALICO Journal*, 34.2, 243-258.
- SHARP, H., PREECE, J. & ROGERS, Y. 2019. *Interaction Design: Beyond Human-Computer Interaction*, West Sussex, Wiley & Sons.
- SMITH, L. T. 2012. *Decolonizing methodologies: research and indigenous peoples, second edition*, Zed Books.
- SORO, A., BRERETON, M., TAYLOR, J. L., HONG, A. L. & ROE, P. 2016. Cross-Cultural Dialogical Probes. *Proceedings of the First African Conference on Human Computer Interaction*. Association for Computing Machinery.

- SZAPIRO, D., KUTAY, C., GARCIA, J., RAFFE, W. & GREEN, R. Learning on Country A Game-Based Experience of an Australian Aboriginal Language. *Proceedings of the International Conference of Innovation in Media and Visual Design (IMDES 2020)*, 2020. Atlantis Press, 60-68.
- TAYLOR, J. L. 2020. *Ngana Wubulku Junkurr-Jiku Balkaway-Ka: The intergenerational co-design of a tangible technology to keep active use of the Kuku Yalanji Aboriginal language strong*. Queensland University of Technology.
- TAYLOR, J. L., ABORIGINAL SHIRE COUNCIL, W. W., SORO, A., ESTEBAN, M., VALLINO, A., ROE, P. & BRERETON, M. 2020. Crocodile Language Friend: Tangibles to Foster Children's Language Use. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- TAYLOR, J. L., SORO, A., BRERETON, M., HONG, A. L. & ROE, P. 2016. Designing Evaluation beyond Evaluating Design: Measuring Success in Cross-Cultural Projects. *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. Association for Computing Machinery.
- USABILITY.GOV. n.d.,. *User Experience Basics* [Online]. Usability.gov. Available: <https://www.usability.gov/what-and-why/user-experience.html> [Accessed 19 Dec 2022].
- VAVOULA, G. & SHARPLES, M. 2009. Meeting the Challenges in Evaluating Mobile Learning: A 3-Level Evaluation Framework. *International Journal of Mobile and Blended Learning (IJMBL)*, 1, 54-75.
- VITOS, M., ALTENBUCHNER, J., STEVENS, M., CONQUEST, G., LEWIS, J. & HAKLAY, M. 2017. Supporting Collaboration with Non-Literate Forest Communities in the Congo-Basin. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. Association for Computing Machinery.
- W3C. 2018. *WCAG 2.1 at a Glance* [Online]. W3C. Available: <https://www.w3.org/WAI/standards-guidelines/wcag/glance/> [Accessed 19 Dec 2022].
- WARD, M. 2004. The additional uses of CALL in the endangered language context. *ReCALL*, 16.2, 345-359.
- WARD, M. 2018. Qualitative research in less commonly taught and endangered language CALL. *Language Learning*, 22, 116-132.
- YOO, D., DERTHICK, K., GHASSEMIAN, S., HAKIZIMANA, J., GILL, B. & FRIEDMAN, B. 2016. Multi-Lifespan Design Thinking: Two Methods and a Case Study with the Rwandan Diaspora. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- ZAMAN, T., WINSCHIERS-THEOPHILUS, H., YEO, A. W., TING, L. C. & JENGAN, G. 2015. Reviving an Indigenous Rainforest Sign Language: Digital Oroo' Adventure Game. *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*. Association for Computing Machinery.