A catalogue of evidence for transport interventions

Design and initial prototype

July 2024

Vivienne Ivory, WSP Research & Innovation, Wellington
Edgar Pacheco, WSP Research & Innovation, Wellington
Francene Thomas, WSP Research & Innovation, Wellington
Daniel Cooper, WSP Research & Innovation, Wellington
Louise Malcom, WSP Research & Innovation, Wellington
Suzanne Nugent, WSP, United Kingdom
Joel Burton, WSP Research & Innovation, Wellington

Keywords: Catalogue; design; evidence; interventions; prototype; transport

An important note for the reader

NZ Transport Agency Waka Kotahi (NZTA) is a Crown entity established under the Land Transport Management Act 2003. The objective of NZTA is to undertake its functions in a way that contributes to an efficient, effective and safe land transport system in the public interest. Each year, NZTA funds innovative and relevant research that contributes to this objective.

The views expressed in research reports are the outcomes of the independent research and should not be regarded as being the opinion or responsibility of NZTA. The material contained in the reports should not be construed in any way as policy adopted by NZTA or indeed any agency of the New Zealand Government. The reports may, however, be used by New Zealand Government agencies as a reference in the development of policy.

While research reports are believed to be correct at the time of their preparation[1], NZTA and agents involved in their preparation and publication do not accept any liability for use of the research. People using the research, whether directly or indirectly, should apply and rely on their own skill and judgement. They should not rely on the contents of the research reports in isolation from other sources of advice and information. If necessary, they should seek appropriate legal or other expert advice.

Please note:

This research was conducted under a previous policy context. For example, the research was developed and/or undertaken under the 2021-24 Government Policy Statement for Land Transport. Consequently, references contained in the report may be to policies, legislation and initiatives that have been concluded and/or repealed. Please consider this in your reading of the report and apply your judgement of the applicability of the findings to the current policy context accordingly.

---

[1] This research was conducted April 2023-October 2023.

Acknowledgements

Abbreviations and acronyms

CETI	catalogue of evidence for transport interventions
KonSULT	Knowledgebase on Sustainable Urban Land use and Transport
NZTA	NZ Transport Agency Waka Kotahi
SafetyCube	European Road Safety Decision Support System
SMS	Maryland Scientific Methods Scale
TCAD	Transport Climate Action Directory

Contents

> **Please note:**
>
> This research was conducted under a previous policy context. For example, the research was developed and/or undertaken under the 2021-24 Government Policy Statement for Land Transport. Consequently, references contained in the report may be to policies, legislation and initiatives that have been concluded and/or repealed. Please consider this in your reading of the report and apply your judgement of the applicability of the findings to the current policy context accordingly.

Executive summary

It is difficult to access and use evidence about transport interventions.

Increasingly, transport practitioners and decision-makers want to use evidence to help them decide which transport interventions to use and investments to make – locally, regionally and nationally – across a full range of outcomes in the transport sector. However, it is often challenging to access and use evidence to assess how relevant and effective interventions will be in specific local contexts. These challenges stem from the diversity and complexity of the sources, types and quality of evidence, and the wide range of interventions and outcomes covered by the transport system.

A catalogue of evidence could make it easier to compare and use evidence.

In this report, we present the design for a prototype catalogue of evidence for transport interventions, which aims to overcome these challenges. The design for this catalogue intends to provide a structured way to collate, catalogue and synthesise evidence about transport interventions, which will make it easier for practitioners to compare and use the evidence. The design is based on innovative approaches to evaluating the effectiveness of interventions, examples of other good catalogues and feedback from stakeholders.

We have designed a catalogue that can evaluate interventions in real-world settings.

Our design uses a 'mosaic' approach to evidence. This approach recognises that a wide range of evidence is needed to evaluate the effectiveness of interventions in real-world settings. We followed three steps to develop the design.

Step 1: Developing the requirements for the catalogue of evidence for transport interventions.

To identify the requirements of the catalogue design, we reviewed three existing catalogues: KonSULT (Knowledgebase on Sustainable Urban Land use and Transport), SafetyCube (European Road Safety Decision Support System) and TCAD (Transport Climate Action Directory). We also reviewed user assessments of four catalogues (two transport and two non-transport catalogues), and approaches to 'harmonising' or synthesising evidence. We then held a stakeholder workshop to identify which features in a catalogue decision-makers would find most useful.

From the reviews and stakeholder workshop, we established that the information in a catalogue needs to:
- be comprehensive and diverse
- explain the context that it applies to
- be relevant to New Zealand transport decisions
- enable users to compare evidence about the same intervention function (such as user safety)
- include expert interpretation
- explain the inputs and outputs of interventions.

We also established that the catalogue needs:
- an organising hierarchy
- defined standards
- comparison, search and gap-analysis features
- documentation
- consistent language

- a user-friendly interface.

Step 2: Designing and reviewing a prototype catalogue of evidence for transport interventions.

The design for the prototype catalogue is grounded in New Zealand's Transport Outcomes Framework.[2] We have designed it to accommodate a wide range of evidence types, interventions and transport outcomes. The catalogue design links individual sources of evidence, hierarchically, to the 5 outcomes in the Transport Outcomes Framework and the 12 benefit clusters in the Land Transport Benefits Framework.[3]. Collectively, these represent an intervention's intended functional benefit (such as changes to user safety or resource efficiency). This hierarchy, evidence standards (standards about which types of evidence are included) and metadata standards (standards about which data fields are included) form the catalogue's framework. The hierarchy ensures that the catalogue presents diverse sources of evidence in a standardised way.

We built a digital prototype of the catalogue, to demonstrate the design features. To do this, we used a small selection of studies to extract information that we could use to filter evidence. This information includes:

- source details (such as the title, abstract and publication type)
- intervention details (such as the type, cost, location, geographical scale, mode and urbanicity of the intervention; observed changes following the intervention; and relevance of the intervention to New Zealand urban tiers)
- evaluation-quality details (such as the study design, assessment type and type of data used, and whether the evaluation was peer reviewed)
- details of how the intervention relates to the Transport Outcomes Framework and benefits clusters.

Users can search for, and compare, information in the catalogue using filters such as intervention type, intervention outcome and relevance of the intervention to New Zealand. The catalogue design includes summaries and syntheses of the evidence, as well as information on costs, benefits and transferability. Users can access the source documents, view the results as a list or on a map, and export or share the results.

We developed a qualitative harmonisation heuristic process – or 'rule of thumb' decision-support tool − based on the realist-review method. This heuristic process helps users form a judgement when they compare dissimilar evidence or consider whether an intervention is transferrable to a different context. It encourages them to transparently incorporate their supplementary knowledge when they assess the likelihood that an intervention will be effective in a novel setting.

Step 3: Agreeing the scope of the catalogue of evidence for transport interventions with the steering group

In the final step of this project, the project steering group reviewed the design and prototype against the requirements we established in Step 1. With the steering group, we considered whether the need for a comprehensive, diverse catalogue balances out the effort and resource that is required to implement and maintain it. The steering group recognised that, while it will take more effort and expertise to complete some fields (for example, fields that involve assessing and making a judgement on an intervention's relevance to New Zealand), the value of the catalogue lies in being able to synthesise this data for users.

---

[2] For more information, visit www.transport.govt.nz/area-of-interest/strategy-and-direction/transport-outcomes-framework

[3] For more information, visit www.nzta.govt.nz/planning-and-investment/learning-and-resources/benefits-management-guidance/the-land-transport-benefits-framework/introduction-to-the-benefits-framework

Transport practitioners think the catalogue is valuable and worth the effort needed to develop and maintain it.

Transport practitioners we consulted saw the prototype catalogue as a valuable tool for making evidence-based decisions to improve New Zealand's transport system. We conclude that there is a need and a demand for the catalogue, but we also acknowledge the challenges involved, and resources required, to establish and maintain it. This report provides recommendations for further developing and implementing the catalogue, taking into account feedback we received from users and the steering group.

Abstract

Increasingly, transport practitioners and decision-makers want to use evidence to help them decide which transport interventions to use and investments to make – locally, regionally and nationally – across the full range of transport outcomes. However, it is often challenging to access and use evidence to assess how relevant and effective interventions will be in specific local contexts. These challenges stem from the diversity and complexity of the sources, types and quality of evidence, and the wide range of interventions and outcomes covered by the transport system.

This report presents the design for a prototype catalogue of evidence for transport interventions, which aims to overcome these challenges. The design for this catalogue intends to provide a structured and harmonised way to collate and catalogue evidence about transport interventions, which will make it easier for practitioners to compare and use the evidence.

The design is based on examples of other good catalogues and feedback from stakeholders, which are incorporated into the prototype alongside evidence and metadata standards. The design uses a 'mosaic' approach to evidence. This approach recognises that a wide range of evidence is needed to evaluate the effectiveness of interventions in real-world settings.

A digital prototype was built to demonstrate the design features, which include filters for extracting, assessing and synthesising information from individual sources. A qualitative harmonisation heuristic process was developed to help users harmonise evidence from dissimilar contexts and study types, to transparently determine the likelihood that an intervention will be effective in a novel application.

While ambitious, a comprehensive, diverse catalogue that incorporates expert synthesis will add value to existing tools available to transport practitioners. With engagement from users, the catalogue will help improve the generation of evidence about transport interventions, by setting standards for its evaluation and reporting that will, ultimately, help address information gaps.

1 Introduction

Increasingly, transport practitioners and decision-makers want to use evidence on the effectiveness of transport interventions to help them decide which ones to use and invest in – locally, regionally and nationally − across the full range of transport outcomes.[4] However, it is often challenging to access and use evidence.

NZ Transport Agency Waka Kotahi (NZTA) has identified the need for a structured way to collate and catalogue evidence about transport interventions. A well-designed catalogue of evidence for transport interventions (CETI) will increase practitioners' capacity to efficiently assess and compare, and ultimately select, appropriate transport interventions.

Evidence comes in many forms, but the robustness and quality of the evidence, and its coverage of outcomes and interventions, varies considerably. The design for a CETI needs to recognise that:

- evidence is needed across the full range of transport outcomes, to monitor and evaluate previous and current investments, and plan for and prioritise new investments
- evidence comes from many different sources, ranging from a systematic literature review published behind an academic paywall to a peer reviewed NZTA research report or case study referenced on a council website
- evidence can stem from evaluations that use different types of data and methodologies
- the transport system uses a broad range of interventions, including policy and regulations, infrastructure, behavioural measures (such as financial incentives to use public transport) and urban environmental changes (such as urban redevelopment initiatives).

For a CETI to be useful to transport practitioners, it must be able to compare diverse evidence across a range of transport outcomes. Rather than simply collating evidence, a useful catalogue would harmonise diverse evidence, so that transport practitioners can make sense of a body of work, consider how it relates to their context, and be confident about what they infer from it.

This report describes how we developed, during 2023, the design for a prototype CETI that responds to New Zealand's needs.

1.1 Purpose of this research

NZTA wants to design a CETI that structures evidence, so that interventions can be examined by their:

- applicability
- cost (operational and whole-of-life)
- responsiveness to differing local needs.

The research goal is a catalogue of structured and searchable information about transport interventions.

The research objective is to develop the design for a structured, searchable CETI, by:

- developing a framework for cataloguing evidence on the effectiveness, transferability and cost of transport interventions in a structured way
- developing a standard for transport-intervention metadata, so that evidence is machine searchable
- developing standards about which evidence is included in the catalogue, so that the content is consistent

---

[4] For more information about transport outcomes, visit www.transport.govt.nz/area-of-interest/strategy-and-direction/transport-outcomes-framework

- developing methods to harmonise results, so users can make inferences about significant contextual factors
- testing a user-friendly prototype of the catalogue, using case studies and example literature reviews.

1.2 Approaches to using evidence about transport interventions

Our approach to designing the catalogue was guided by lessons learnt from previous literature reviews and catalogue prototypes (Ivory, 2023; Pacheco & Ivory, 2023; Thomas et al., 2022). Based on the lessons we learnt, we know the catalogue needs to be able to support these activities:

- Synthesising evidence.
- Evaluating the effectiveness of interventions in the real world.
- Comparing interventions that have a similar function or purpose.
- Understanding the causes of intervention outcomes.

The traditional hierarchy for evidence on the effectiveness of interventions sets randomised control trials (RCTs) and quasi-experimental study designs as the gold standard. This is partly because they control for other factors that could explain the results (Tannahill, 2008). Unfortunately, there are few examples of this type of evidence when it comes to evaluating transport interventions (Pacheco & Ivory, 2023). Apart from the practicalities of establishing an experimental design when a transport intervention is being implemented and evaluated, Tannahill (2008) makes the case that RCTs are not always appropriate for complex, real-world situations, because it is not ethical or practical to randomly assign study subjects to treatment and control arms.

1.2.1 Synthesising evidence

Systematic reviews and meta-analyses are methods of collating and comparing multiple studies, using set search terms and criteria. In these methods, the results are synthesised to reach a conclusion about the effectiveness or impact of an intervention. To make synthesis feasible, studies are excluded if they do not contain the details required for synthesis across study variables (for example, the duration of an intervention or comparable measures of results). When studies and evaluations relate to similar interventions and use similar evaluation protocols, these necessarily restrictive practices are less of an issue. However, when interventions being synthesised are more complex, or the way they are evaluated is very variable, a systematic approach may not provide useful answers.

1.2.2 Evaluating the effectiveness of interventions in the real world

To evaluate the effectiveness of interventions in real-world public-health settings, Tannahill (2008) proposes a framework that scrutinises evidence in three ways (see Figure 1-1):

1. Ethical principles – *should* we do this? Ethical principles could include a starting point, such as 'do good', or principles, such as equity and sustainability.

2. Methodological strengths of evidence − does the method match the knowledge gap? In-depth, qualitative methods are valuable for understanding *why* something did or did not work; natural experiments provide the opportunity to compare complex interventions; and observational case studies are useful for describing key features, specific applications and outcomes.

3. Theory or pathway − Is there a plausible theory or pathway behind the evidence? Does the underlying theory make sense, considering other knowledge and perspectives? Are there contradicting theories or alternatives that make more sense? Is the evidence an outlier? How coherent is the evidence with the

existing body of evidence? A conceptual pathway between a transport intervention and its intended outcomes can help 'locate' the intervention within the transport system (Mackie et al., 2017).

Figure 1-1   The health-improvement decision-making triangle (reprinted from Tannahill, 2008, p. 387)
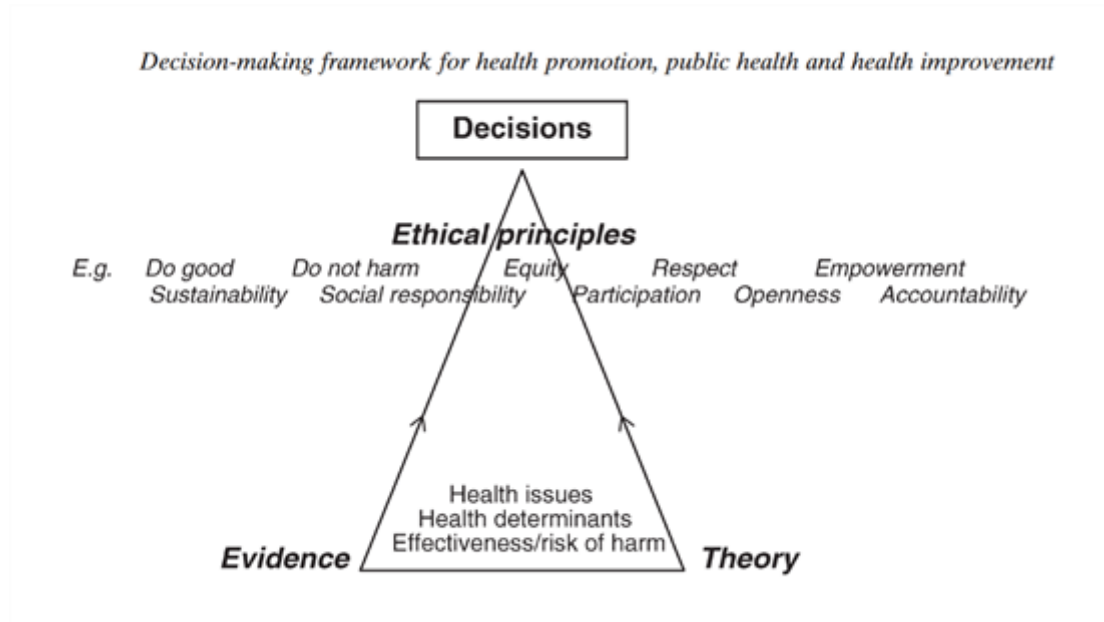
1.2.3   Comparing interventions that have a similar function or purpose

Panter et al. (2019) and Ogilvie et al. (2020) recognise the practical limits to comparing findings from multiple pieces of evidence, particularly the ability to generalise the findings or be confident they will be relevant in different contexts. Rather than focusing on comparing interventions that have a specific type or form, they propose it is more useful to compare interventions based on their function or purpose – what they are intended to achieve.

Panter et al. (2019) argue that a functional approach to comparing interventions across different contexts increases our understanding of *how* interventions work, which is a critical research gap. They also make the case that a mix of evidence types and study designs is needed to assess the magnitude and plausibility of causal mechanisms. They propose using a 'mosaic' approach to building an evidence base. This approach integrates different forms of evidence to build a bigger picture of effectiveness.

1.2.4   Understanding the causes of intervention outcomes

If we use evidence to evaluate the effectiveness of interventions, it needs to help us think about what causes outcomes, or how the variables in an evaluation address causal factors and pathways. Rees et al. (2020) pose a challenge to those who generate evidence. They say that evidence should be about increasing certainty and knowledge:

> *The challenge becomes one of providing a logical argument, supported by sufficient evidence, that the intervention made an important contribution to the observed results. This requires that:*
>
> a.   *the intervention has a clear, explicit theory of change*
>
> b.   *the intervention is implemented as set out in the theory of change, or any changes are captured in an updated theory of change*

c. *the theory of change is supported by, and confirmed by, evidence of observed results and underlying assumptions*

d. *other influencing factors have been assessed and either shown not to have made a significant contribution, or their relative role in contributing to the desired result has been recognised. (Rees et al., 2020, p. 35)*

1.3 Principles for designing a catalogue of evidence for transport interventions

Drawing on these lessons from published literature, we based the design of the CETI on these principles (see Figure 1-2).

- Build an 'evidence mosaic' that covers a diverse range of outcomes, interventions and assessment types.
- Include comprehensive information, to enable users to make sense of the evidence and understand the causes of outcomes.
- Focus on the function of interventions – what they are seeking to change.

Figure 1-2   Principles for designing a catalogue of evidence for transport interventions
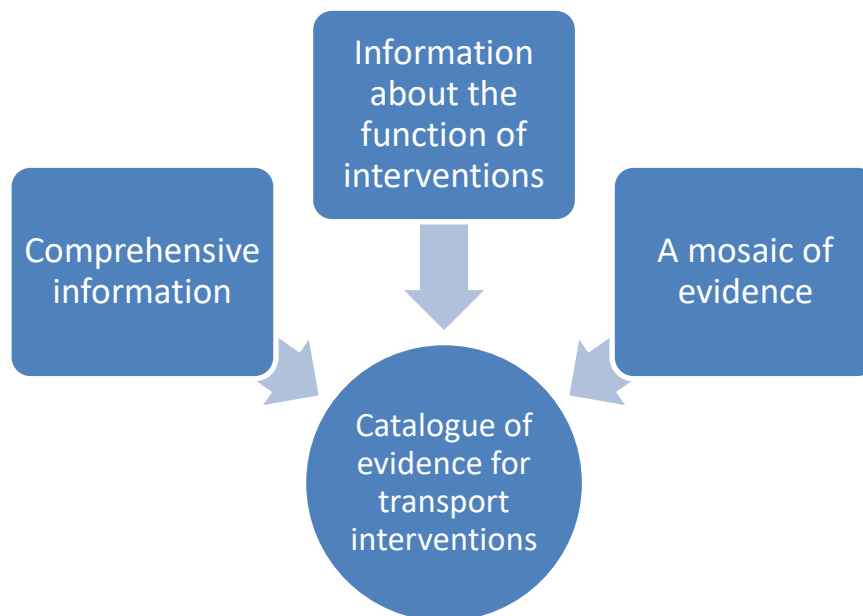
1.4 Process for designing a catalogue of evidence for transport interventions

We followed three steps to develop the design for the CETI:

1. Developing the requirements for the CETI (see section 2).
2. Designing and reviewing a prototype CETI (see section 3).
3. Agreeing the scope of the CETI with the steering group (see section 4).

This report describes the project's three steps, and then presents our recommendations and conclusions.

2 Step 1: Developing the requirements for the catalogue of evidence

The first step in the project was understanding what a 'good' catalogue would look like. This involved asking two questions:

1. What information is needed to evaluate the effectiveness of interventions?
2. What form does the information need to be in?

We sought answers to these questions by:

- reviewing existing catalogues
- reviewing best practice for cataloguing and managing evidence
- holding a workshop with stakeholders.

2.1 Desktop review of existing catalogues

To our knowledge, there are currently no publicly accessible evidence catalogues that include the full range of outcomes and interventions that NZTA wants to include. However, we can gain some useful lessons from the more specialised tools that are available, to help us develop the requirements for the CETI.

2.1.1 Existing tools and approaches

We reviewed three existing intervention catalogues:

- The Knowledgebase on Sustainable Urban Land use and Transport (KonSULT).[5]
- SafetyCube (the European Road Safety Decision Support System).[6]
- Transport Climate Action Directory (TCAD).[7]

For each catalogue, we examined information available about its purpose, audience, taxonomy, assessment types, search features, cost and user assistance (such as help functions), and its arrangements for updates, maintenance, quality control and evaluation.

We chose these catalogues because they cover specific transport-related areas that are relevant to this project and offer useful guidance. Each catalogue has a different focus: sustainable urban transport (KonSULT), transport safety (SafetyCube) and transport decarbonisation (TCAD). However, they all exist to support decision-makers and policymakers, by organising and synthesising information from different sources. Most of the information they contain is based on research and has been peer reviewed.

All three catalogues are designed to be online tools and available to wide audiences – features that align with the purpose of this project. We identified, but excluded, other intervention catalogues,[8] either because their format is not interactive and user-centred, or because their taxonomy and structure does not meet the needs of this project.

See Table 2-1 for a summary and comparison of the features of the three catalogues.

---

[5] For more information, visit www.konsult.leeds.ac.uk

[6] For more information, visit www.roadsafety-dss.eu/#

[7] For more information, visit www.itf-oecd.org/transport-climate-action-directory-measures

[8] These catalogues include the health-related Catalogue of interventions addressing vaccine hesitancy and the education-related What Works Clearinghouse.

2.1.1.1   Knowledgebase on Sustainable Urban Land use and Transport

KonSULT is an online tool, designed to help policymakers, transport and urban specialists[9] and interest groups understand and achieve sustainable urban transport. The tool helps users identify appropriate policy measures and packages (see Figure 2-1).

KonSULT was available from 2001 but was officially launched in 2002 at the first workshop of the World Conference on Transport Research Society Special Interest Group on Urban Transport Policy in Leeds, UK. It ran until 2016.

The project was funded by the European Commission, the UK Department for Transport, the UK Engineering and Physical Sciences Research Council and the Rees Jeffreys Road Fund. The latest version of KonSULT was funded by the European Commission's CH4LLENGE project.

Due to funding cuts and restrictions related to the COVID-19 pandemic, KonSULT is no longer being updated. No information is available about the cost to develop or maintain KonSULT.

Figure 2-1   KonSULT home page (www.konsult.leeds.ac.uk)
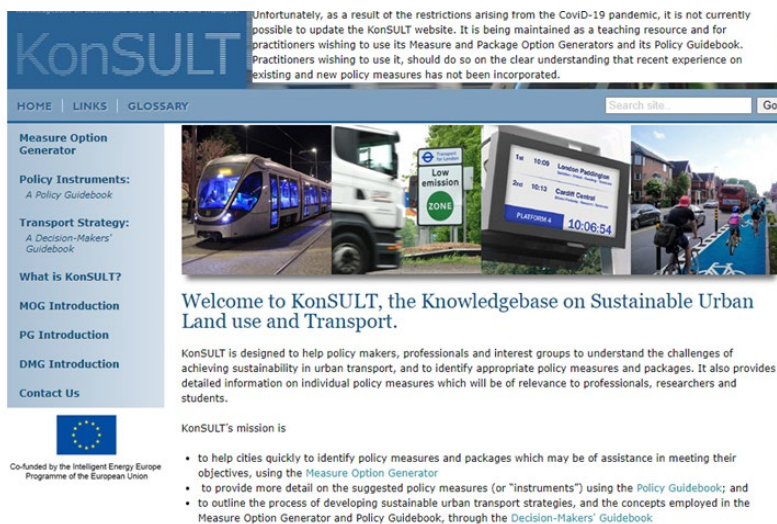
Our assessment of KonSULT

KonSULT is a collection of policy measures that focus on urban-transport policy. The tool's purpose, audience and requirements are clearly defined and consistently communicated on different pages of the KonSULT website. The tool can be accessed from anywhere, but it is targeted at a European audience. KonSULT only supports English, which may pose a challenge to non-English speakers.

KonSULT was developed using 64 policy measures. The measures are grouped into six categories:

- Land use.
- Infrastructure.
- Management and service.
- Attitudinal and behavioural.
- Information provision.

---

[9] This term includes specialists who focus on non-transport urban matters, such as urban design or land-use planning.

- Pricing.

KonSULT has a Measure Option Generator feature, which allows users to search for policy measures by objectives, problems or indicators. The search generates a list of measures, including their policy ranking, cost (high, medium, low or neutral) and timescale (short, medium or long).

Each policy measure has easy-to-read content, divided into the following short sections:

- Summary.

- Taxonomy and description.

- First-principles assessment.

- Evidence on performance.

- Policy contribution.

- References.

The policy measures are based on gathered and evaluated resources, and include cross-sectional studies, case reports, reviews and expert opinion. An editorial team provided a quality-assurance function, by establishing and applying a protocol for assessing every source of information that is included in the catalogue (May & Taylor, 2002) The sources are cited in the sections that describe each policy measure, and are listed in the references section. However, the website does not explain which criteria are used to identify and assess the sources that inform the policy measure. Also, while these sources are properly cited, KonSULT does not update the links to retrieve them, so we could not find some of the older ones.

The policy contribution section is an interesting feature; it provides an assessment of the policy measure's impact. Another interesting feature is KonSULT's ranking system, which assesses different policy aspects, across and within policy measures. However, the rationale and process for the ranking system are not explained.

The KonSULT website provides documents and training materials in text format, but it does not provide multimedia formats such as video.

After KonSULT was launched, its project team evaluated how effective it was proving for transport and urban specialists who were using the tool (May et al., 2018). The evaluation found that the catalogue was most useful to those with less experience. Since the project finished in 2016, the KonSULT tool remains accessible, but is no longer being updated. This means that new measures and recent research are not included in the tool.

2.1.1.2   SafetyCube

SafetyCube provides European and international evidence on what causes road crashes and non-fatal serious injuries, and evidence on interventions that have been shown to mitigate these threats (see Figure 2-2). SafetyCube is targeted at road and transport policymakers. It aims to help them select the most appropriate measures to reduce road-user casualties and the severity of crashes. The tool contains information on a wide variety of behaviour-, infrastructure- and vehicle-related risk factors, and the effectiveness and cost-effectiveness of measures that can reduce these risk factors.

SafetyCube was developed over three years (2015−2018) and has been operating since April 2017. It cost €5,790,111 to set up SafetyCube; which was funded by the European Commission under Horizons 2020 (the EU Framework Programme for Research and Innovation) road safety domain.

Figure 2-2   SafetyCube home page (www.roadsafety-dss.eu/#/knowledge)

Our assessment of SafetyCube

The SafetyCube website clearly explains the tool's objectives and target field, which is road safety. It has a user-friendly interface that allows users to easily locate evidence by searching by keyword, risk factor, safety measure, road-user group and accident category.

SafetyCube is targeted at public authorities, policymakers, media outlets, non-government organisations and researchers worldwide, although it focuses on Europe. It consists of 34 risk factors and 50 safety measures, which are grouped into four categories:

- Behaviour.
- Infrastructure.
- Vehicle.
- Post-impact care.

When a user selects a measure, they can filter their search results by safety measure, road-user group, road type and country. The search results provide a list of approaches related to the selected measure, which are graded for effectiveness using a traffic-light rating.

SafetyCube uses quantitative evidence to develop the safety measures. This includes before−after studies, cross-sectional studies, induced exposure studies, time-series studies, and different types of statistical methods (simple comparisons of counts or means, regression analyses, empirical Bayes and hazard rates). It gives a synopsis for each safety measure, which summarises and discusses the research related to it. (These synopses are more detailed than those in KonSULT.) The website's 'Study Page' gives more information about each study that informs the safety measure. The metadata behind each study gives descriptive information commonly used by research libraries (such as title, author, year, abstract and keywords), as well as information about the study's limitations and methodology, and the impact of its results on road safety.

An additional feature of SafetyCube is the Economic Efficiency Evaluation (E3) calculator, which combines evidence about a safety measure's effectiveness with its costs. This includes information about crash costs in European countries, which, according to the project team, allows users to express a safety measure's costs and benefits in monetary values for cost-benefit-analysis purposes.

SafetyCube was developed by consulting with stakeholders and experts. The project team identified, assessed and coded the evidence and developed the synopses. In addition, a pool of independent experts

reviewed the content and analysis, using predefined quality criteria, and a native English speaker performed a language check.

The SafetyCube website contains guidance on how to use the tool, including links to short videos on a YouTube channel. Although the project and its methodology are documented, we are not aware that it has been formally evaluated.

2.1.1.3   Transport Climate Action Directory

TCAD is an online public database of transport policy measures to reduce CO_2 emissions (see Figure 2-3). It contains 80 mitigation measures alongside evidence to assess their effectiveness. TCAD aims to help decision-makers translate their decarbonisation goals into actions, so they achieve their climate objectives. It gives them a range of options to achieve transport-related decarbonisation outcomes in their specific context.

TCAD was launched by the International Transport Forum in July 2020, to support countries making their first revision of their Nationally Determined Contributions for the 2021 Conference of Parties (COP26) of the United Nations Framework Convention on Climate Change. We did not find any information on how TCAD is funded.

Figure 2-3   TCAD home page (www.itf-oecd.org/transport-climate-action-directory-measures)
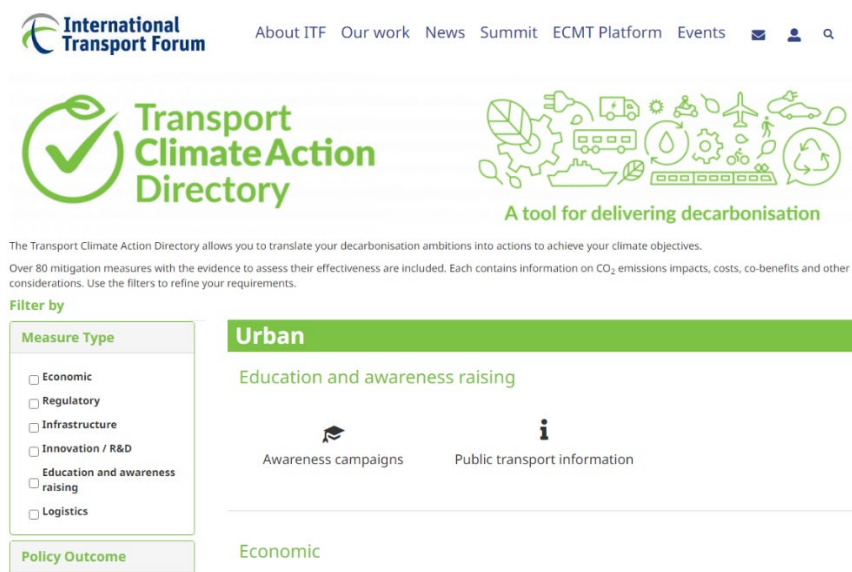
Our assessment of TCAD

TCAD is targeted at decision-makers and policymakers, worldwide, who are involved in climate and transport policy. The tool is current, and the information underpinning it is updated periodically.

TCAD covers decarbonisation of the whole transport system. It contains 80 evidence-based measures to mitigate CO_2 emissions. The measures are grouped into five policy outcomes:

- Improved design, operations and planning of transport systems.
- Electrification.
- Low-carbon fuels and energy vectors.
- Mode shift and demand management.
- Innovation and up-scaling.

Experts from the International Transport Forum analysed and wrote up the measures, and their work was peer reviewed by external specialist researchers. In contrast with SafetyCube, the evidence that informs the measures in TCAD is not restricted to quantitative material. It includes evidence from modelling, experiments and quasi-experiments, strategy compilations and policy reviews.

For each measure, there is a broad description of the potential costs involved, including estimates in some cases. However, unlike KonSULT, TCAD does not provide a scale of costs. The tool also offers brief summaries of the impact that measures have on reducing CO_2 emissions, as well as information on their co-benefits and contextual considerations.

Users can share the results of TCAD searches on social media or use the online submission system to suggest new measures or updates to existing ones.

We are not aware that TCAD has been formally evaluated.

Table 2-1   Summary and comparison of three online catalogues of transport interventions

	KonSULT	SafetyCube	TCAD
Purpose	To help users understand the challenges of achieving sustainable urban transport, and identify appropriate policy measures and packages	To help users make road-safety decisions and choose appropriate strategies, measures and cost-effective approaches, to reduce casualties from road accidents	To help users translate their decarbonisation goals into actions, to achieve their climate objectives
Area of interest	Europe	Worldwide, especially Europe	Worldwide
Field	Urban-transport policy	Road safety	Transport policy related to mitigating the effects of climate change
Target audience	Policymakers, transport and urban specialists and interest groups	Public authorities, industries, research institutes, mass media and non-government organisations	Policymakers
Update and maintenance	Last updated in 2016	Project development ended in 2018. No information is available on how often the tool is updated	Updated periodically
Language	English	English	English
Taxonomy	64 policy measures, grouped into 6 categories: • Land use • Infrastructure • Management and service • Attitudinal and behavioural • Information • Pricing	34 risk factors and 50 safety measures, grouped into 4 parts: • Behaviour • Infrastructure • Vehicle • Post-impact care (measures only)	80 CO_2 mitigation measures, grouped into 5 policy outcomes: • Improved design, operations and planning of transport systems • Electrification • Low-carbon fuels and energy vectors • Mode shift and demand management • Innovation and up-scaling

19

	KonSULT	SafetyCube	TCAD
Quality assurance	Data quality was assessed by an editorial team that used a quality-assurance protocol	Risk factors and measures are developed, analysed and summarised by the project team and stakeholders. This work is peer reviewed by independent experts who check the content analysis and content against predefined criteria. A native English speaker performs a language check	Mitigation measures are written by experts from the International Transport Forum and peer reviewed by external specialist researchers
Cost information	Yes. Each measure was assigned a cost category (high, medium, low or neutral)	Yes. The tool includes the Economic Efficiency Evaluation (E3) calculator, which combines evidence about a measure's effectiveness with its costs	Yes. The description for each measure includes its potential costs, with estimates in some cases. It does not provide a scale of costs
Assessment types	- Cross-sectional studies - Case reports - Reviews of policy measures - Expert opinion	- Before–after studies - Cross-sectional studies - Case-control studies - Induced exposure studies - Time-series studies - Statistical methods (simple comparisons of counts or means, regression analyses, empirical Bayes and hazard rates)	- Modelling - Experiments - Quasi-experiments - Quantitative Q studies - Sensitivity analysis - Compilation of strategies - Reviews of policy measures
Features	A search feature for policy measures that uses dropdown menus, a keyword search, hyperlinks from a complementary instruments section or a filter. The filter is the basis for the Measure Option Generator	A stand-alone calculator (Economic Efficiency Evaluation or E3) that combines evidence about a measure's effectiveness with its costs	Brief summaries of the impact of measures on CO_2 emissions, the co-benefits of measures and contextual considerations Search results that can be shared on social media An online submission system for suggesting new measures or updates to existing ones
User support	Yes. Text-based guidelines on how to use the tool	Yes. Links to downloadable guidelines and multimedia content	Yes. Video-based tutorials. Tool is also designed to be intuitive
Evaluation	Yes. The tool was self-evaluated by the project team and transport and urban specialists who used the tool. It has not been independently evaluated	Unknown	Unknown

2.1.2 Evaluation of the interface of existing catalogues

A catalogue's success hinges on the useability of its interface. We evaluated the user interfaces of four existing catalogues to identify which qualities we wanted to use in the design of the CETI.

2.1.2.1 Principles for designing a catalogue's user interface

Shneiderman (1998) identifies eight principles to apply when designing a user interface:

1. **Strive for consistency**: Users should not have to wonder whether different words, situations or actions mean the same thing: Use identical terminology in prompts, menus and help screens, and consistent commands throughout the interface.

2. **Enable frequent users to use shortcuts**: Hide shortcuts from novice users, but make them available to frequent or expert users, to speed up their interaction with the material.

3. **Offer informative feedback**: The interface should keep users informed about what is going on as they engage with the site, by giving them appropriate, timely feedback.

4. **Design dialogue that enables users to easily reach an endpoint**: Interfaces should not contain irrelevant information, because every extra unit of information in an interface competes with relevant units of information. The design should speak users' language; it should use words, phrases and concepts that users are familiar with, rather than internal jargon. Ideally, it will be unnecessary to explain to users how to use the interface; however, they may sometimes need documentation to help them complete tasks.

5. **Prevent errors, or handle them simply**: Good error messages are important, but the best designs carefully prevent problems occurring in the first place. When error messages are required, they should be expressed in plain language (no error codes), precisely indicate the problem and suggest a constructive solution.

6. **Permit users to reverse their actions**: Users often perform actions by mistake. They need a clearly marked 'emergency exit' to leave the unwanted action.

7. **Support users to feel in control**: Design the interface so that users initiate actions rather than only respond to information that is provided, so they feel they are in control.

8. **Cut down how much users need to remember**: Make the site's elements, actions and options visible to users, and avoid them having to remember information they entered on previous screens.

We evaluated four catalogues of evidence according to these principles, by using eleven questions to interrogate the catalogues from a user's perspective. We used the questions to rank the catalogues' useability from 1 (less useable) to 3 (more useable). We evaluated transport catalogues (KonSULT and SafetyCube) and non-transport catalogues (What Works Clearinghouse[10] and Substance Abuse and Mental Health Services Administration's Evidence-Based Practices Resources Center[11]), with varying levels of complexity.

Overall, we find that, with a good design, a catalogue of evidence can adhere to these useability principles. We rated the simplest catalogue (Evidence-Based Practices Resources Center) the highest for useability. It rated poorly for only one principle – enabling frequent or expert users to use shortcuts. However, in some cases, its high useability rating was due to its relative simplicity.

We find that SafetyCube's features and functions are more useable than KonSULT's. This is largely because the KonSULT interface allows users to make errors without giving them a warning or the means to readily

---

[10] For more information, visit https://ies.ed.gov/ncee/wwc

[11] For more information, visit www.samhsa.gov/resource-search/ebp

correct their mistake. For example, users can select multiple checkmarks in a dropdown menu when the tool can only run one at a time. We also find the language in KonSULT is less useable than that in other catalogues, because it is more technical.

The results of our evaluation of each catalogue are summarised in Table 2-2 (KonSULT), Table 2-3 (SafetyCube), Table 2-4 (What Works Clearinghouse) and Table 2-5 (Evidence-Based Practices Resources Center).

Table 2-2   Evaluation of the useability of KonSULT user interface

Principle	Question	Evaluation rating	Comments
Strive for consistency: Standardising how information is presented	Are the terms, menus and help screens consistent?	2	KonSULT mainly uses dropdown menus, but also has checkboxes for choosing objectives, problems and indicators. The checkboxes cause a problem, because users can select multiple options, but the tool can only run one at a time
Strive for consistency: Standardising how information is presented	Is the content in a logical location?	2	The 'previous page' button is above the 'continue' button and can be clicked on accidently when users want to continue
Enable frequent or expert users to use shortcuts: Providing quicker methods to complete tasks	Is the tool made with the needs of experienced users in mind?	2	KonSULT does not seem to offer more experienced users a way to move through the system quicker If KonSULT had every option on the same page, users would not need to go through previous pages to adjust settings A potential option to improve useability would be allowing experienced users to adjust, copy or paste code into a task window
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Does the user know where they are in a process? Is the information presented in a way that is easy to digest?	2	Unless users read the full list of instructions before using KonSULT, they will not know how many steps are in the tool If all the search information was on one page, users would not need to remember information from previous pages to interpret their results table. It would also be useful if users could clear (reset) the options they select on each page
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Is all the information relevant and easy to understand?	3	The paragraphs are short and do not contain unnecessary text. Additional information is available on other pages if it is needed The interface could make the information easy for newer users to understand, by using features such as icons

Principle	Question	Evaluation rating	Comments
Design dialogue that enables users to easily reach an endpoint: Ensuring users know what the result of their action will be	Does the user know what they have done after performing each action?	2	KonSULT does not update users. It retrospectively informs them that they have made a mistake, rather than reducing their ability to make mistakes
Prevent errors, or handle them simply: Designing systems to reduce errors and giving users clear and easy instructions	Is feedback communicated effectively to the user?	1	The feedback and error messages are not obvious, which makes the tool difficult for new users
Permit users to reverse their actions	Are errors reduced where possible, and are they handled appropriately?	1	When the user selects the 'start again' button, it takes them back to the first page but does not clear their selected options. This makes it easier for them to make an error, as they need to manually change the options they previously selected
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the system make the user feel in control of their actions?	1	Users can encounter a number of error messages, which may make them feel less in control of the tool. KonSULT could also be made easier to navigate and have fewer pages, to support users' flow through the interface
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the tool use words, phrases and concepts that users are familiar with?	2	Newer users may find KonSULT's technical language and results difficult to interpret, although further information is available on another page. It may be useful to provide an example of how to read the results
Cut down how much users need to remember: Not requiring users to recall lots of information to use the tool	Does the tool make it easy for the user by reducing the number of things they need to remember?	1	Users have no reference to what they selected on the previous page. No effort has been made to reduce the memory load for users

Table 2-3   Evaluation of the useability of SafetyCube

Principle	Question	Evaluation rating	Comments
Strive for consistency: Standardising how information is presented	Are the terms, menus and help screens consistent?	3	SafetyCube has five entry points, which are located at the top of the screen. Users can click between them and see the options for each entry point
Strive for consistency: Standardising how information is presented	Is the content in a logical location?	3	The location of options at the top and side of the screen is a user-friendly feature
Enable frequent or expert users to use shortcuts: Providing quicker methods to complete tasks	Is the tool made with the needs of experienced users in mind?	1	SafetyCube does not have features that show it was made with the needs of experienced users in mind

Principle	Question	Evaluation rating	Comments
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Does the user know where they are in a process? Is the information presented in a way that is easy to digest?	3	Users provide information on the main screen and are then taken to the responses, which include a colour-coded effectiveness rating and synopsis for each measure
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Is all the information relevant and easy to understand?	3	The tool has many resources available to help users, which are often in video format. These resources are easily accessible. Having them on a separate page reduces how much information is presented on the interface
Design dialogue that enables users to easily reach an endpoint: Ensuring users know what the result of their action will be	Does the user know what they have done after performing each action?	3	As users select options, the tool updates. This can help users understand what each option is doing
Prevent errors, or handle them simply: Designing systems to reduce errors and giving users clear and easy instructions	Is feedback communicated effectively to the user?	2	The keyword search gives you options as you type. However, it appears that users cannot search for a term that does not auto populate
Permit users to reverse their actions	Are errors reduced where possible, and are they handled appropriately?	3	SafetyCube is designed to reduce user error
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the system make the user feel in control of their actions?	3	As users select options, the tool updates. This can help users understand what each option is doing
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the tool use words, phrases and concepts that users are familiar with?	3	The tool uses plain language and does not contain difficult terminology
Cut down how much users need to remember: Not requiring users to recall lots of information to use the tool	Does the tool make it easy for the user by reducing the number of things they need to remember?	1	SafetyCube does not have features that reduce the number of things that users need to remember

Table 2-4   Evaluation of the useability of What Works Clearinghouse

Principle	Question	Evaluation rating	Comments
Strive for consistency: Standardising how information is presented	Are the terms, menus and help screens consistent?	2	The open question box for keywords and the dropdown for topics allow users to select multiple options; however, the 'grade band' dropdown menu does not. This means users must perform a separate search for

Principle	Question	Evaluation rating	Comments
			each grade band, or exclude the filter to see results for all grade bands Some buttons have images of resources, but they are not very obvious Once the tool has filtered search results, users have additional options
Strive for consistency: Standardising how information is presented	Is the content in a logical location?	3	The tool functions are well positioned
Enable frequent or expert users to use shortcuts: Providing quicker methods to complete tasks	Is the tool made with the needs of experienced users in mind?	1	What Works Clearinghouse does not have features that enable experienced users to use shortcuts
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Does the user know where they are in a process? Is the information presented in a way that is easy to digest?	3	The tool functions are well positioned What Works Clearinghouse has videos available to help users understand the ratings and tiers it assigns to evidence. The colours and wording used to explain evidence tiers makes it easy for users to understand them
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Is all the information relevant and easy to understand?	3	What Works Clearinghouse has limited information on how to use it, but this is not needed as the tool is easy to use. It is useful that users can see the abstract for each article, and that each article's title is linked to a page containing the full article
Design dialogue that enables users to easily reach an endpoint: Ensuring users know what the result of their action will be	Does the user know what they have done after performing each action?	3	Yes. The tool is very user friendly
Prevent errors, or handle them simply: Designing systems to reduce errors and giving users clear and easy instructions	Is feedback communicated effectively to the user?	3	There are few options for users to make errors, so there is little opportunity for the tool to give feedback
Permit users to reverse their actions	Are errors reduced where possible, and are they handled appropriately?	3	The tool reduces user errors very well
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the system make the user feel in control of their actions?	3	The search-results table updates to respond to a user's choices. However, it is not clear that users can reorder the table's column titles by clicking on them. This could be made clearer
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the tool use words, phrases and concepts that users are familiar with?	3	What Works Clearinghouse uses unfamiliar terms like 'evidence tier', but alongside common words, like 'strong' and 'promising', it is easy for users to interpret. The website also

Principle	Question	Evaluation rating	Comments
			contains resources to explain the ratings it uses
Cut down how much users need to remember: Not requiring users to recall lots of information to use the tool	Does the tool make it easy for the user by reducing the number of things they need to remember?	3	A user's selected options are shown alongside the search-results table and can be edited. Each filtered item shows the number of papers related to it

Table 2-5 Evaluation of the useability of Evidence-Based Practices Resources Center

Principle	Question	Evaluation rating	Comments
Strive for consistency: Standardising how information is presented	Are the terms, menus and help screens consistent?	3	The tool uses a standard format seen on other websites and search platforms, which makes it very user friendly Users can search using keywords or filters and use the sorting and viewing tools. The tool displays featured articles before users select them
Strive for consistency: Standardising how information is presented	Is the content in a logical location?	3	The layout is user friendly; it is similar to online shopping platforms
Enable frequent or expert users to use shortcuts: Providing quicker methods to complete tasks	Is the tool made with the needs of experienced users in mind?	1	Evidence-Based Practices Resources Center does not have features that enable experienced users to use shortcuts
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Does the user know where they are in a process? Is the information presented in a way that is easy to digest?	3	Each time a user selects an item the tool's outcomes are updated
Offer informative feedback: Enabling users to know where they are in a process and what is going on with each action	Is all the information relevant and easy to understand?	3	Evidence-Based Practices Resources Center has limited information on how to use it, but this is not needed as the tool is intuitive
Design dialogue that enables users to easily reach an endpoint: Ensuring users know what the result of their action will be	Does the user know what they have done after performing each action?	3	Each time the user performs an action the system is updated. This is slightly inconvenient, as the page reloads with each change, rather than applying all changes at once. However, it makes it clear to users that the change has been applied

Principle	Question	Evaluation rating	Comments
Prevent errors, or handle them simply: Designing systems to reduce errors and giving users clear and easy instructions	Is feedback communicated effectively to the user?	3	Keyword items are displayed as you type. It would be good if they were highlighted as separate keywords, if multiple search terms are required
Permit users to reverse their actions	Are errors reduced where possible, and are they handled appropriately?	3	It is difficult for users to make errors, as the tool is a simple database with basic filters
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the system make the user feel in control of their actions?	3	The tool responds to users, giving them the feeling that they are in control
Support users to feel in control: Giving users the sense that they are in full control of the system	Does the tool use words, phrases and concepts that users are familiar with?	3	Evidence-Based Practices Resources Center does not use difficult terminology. It also groups terms clearly into subheadings
Cut down how much users need to remember: Not requiring users to recall lots of information to use the tool	Does the tool make it easy for the user by reducing the number of things they need to remember?	3	Filters and keywords are visible on the same page as outcomes or results. Each filtered item shows the number of papers related to it

2.2 Review of best practice for cataloguing and managing evidence

2.2.1 Cataloguing evidence

There are useful lessons from the similarities and differences between how the catalogues approach cataloguing. KonSULT, SafetyCube and TCAD each focus on a specific topic, rather than covering a broad range of transport-related subjects. Each has a metadata structure that uses a descriptive approach, including metadata elements such as title, author, date of publication, subject, publisher and description. In each case, the tool uses a set of transport-related measures and a summary of the individual pieces of evidence contained in the catalogue. Although the catalogues are structured and presented in different ways, considerable effort and funding has been invested into each of them, and the summaries, to make them easy to understand and ensure they meet users' needs for policy evidence.

Each catalogue uses a different approach to develop the summaries and score or rank the evidence by relevance and feasibility. The catalogues also approach cost differently. KonSULT assigns potential costs a ranking, while SafetyCube interactively estimates costs, by combining measures and cost-related information.

Based on our assessments of the four catalogues, we identified that best practice in cataloguing evidence includes these activities:

- **Consulting with stakeholders**: This involves engaging with key stakeholders (such as potential users, designers and experts) in the early stages of designing and developing a catalogue, to identify their needs and avoid potential pitfalls.

- **Defining what evidence is**: This involves defining the purpose of the intended catalogue and the types of evidence it will include (such as quantitative and qualitative studies, expert opinion or types of observational research) and giving the rationale for those choices.

- **Developing a taxonomy**: This is an iterative process of building a catalogue's structure. It requires expert input and ongoing reflection and improvement, as evidence is compiled and assessed.

- **Defining quality appraisal**: This involves identifying and explaining how the quality of sources will be graded (such as the strategy, methods and standards that will be used) to decide whether to include them in the catalogue.

- **Peer reviewing selected evidence**: This involves expert academics and practitioners assessing the relevance and quality of the sources, to ensure the catalogue includes high-quality evidence.

- **Preparing content for a lay audience**: This means providing non-expert users with clear, user-friendly summaries and descriptions of the evidence, including any associated typology, and avoiding specialist language or jargon that could confuse them.

- **Updating the catalogue periodically**: This involves adding new, relevant evidence as it becomes available, to keep the catalogue up to date for users. While this is mainly done internally, catalogues could also consider recommendations from external parties.

- **Providing guidance on how to use the catalogue**: This can be done through different means but using multimedia (such as online videos) appears to be most helpful.

- **Enabling content to be shared**: This can include enabling users to share search results by email or social media.

- **Considering other languages**: If the catalogue will be used worldwide, this involves considering what languages will be relevant to users (such as te reo Māori for New Zealand users).

- **Evaluating the catalogue**: This helps ensure the catalogue is fit for purpose.

2.2.2　Harmonising evidence

Catalogues of evidence contain diverse pieces of evidence that need to be made comparable in some way. As discussed in section 2.1, consistency − in language and search functions − is important for a catalogue's ability to make sense of a wide body of evidence. KonSULT, SafetyCube and TCAD all use summaries and, to some degree, use synthesis to harmonise information across pieces of evidence. Other approaches to harmonisation are also possible.

Harmonising data involves combining data from non-identical sources into one useful, cohesive database (Nan et al., 2022). It helps users glean information from otherwise disparate sources. Harmonisation achieves this by standardising, or adjusting, the format of data sources, and the definitions and measurement units they use, so they can be compared. In quantitative studies, harmonisation methods include weighting variables. Harmonisation can also involve making data consistent across individual pieces of evidence in a database, by, for example, setting standards for what data will be included in the database (evidence standards) and how the data will be used (metadata standards). Standards need to be defined for machine learning (Nan et al., 2022).

Not all sources can be readily combined into a cohesive format and cohesive content, where 'like' can be directly compared with 'like'. Rather than restrict data comparisons to similar interventions and methods, Panter et al. (2019) uses a realist-review approach to improve causal explanations across evaluations of diverse interventions that assess the impact that changes to the physical environment have on walking and cycling. They use principles from the realist-review method with qualitative causal-estimation assessment methods to synthesise heterogenous evaluations of interventions. They argue for an approach that compares and synthesises studies based on their function (what an intervention intends to achieve), rather

than an intervention's specific form or context. Using this approach, having a wide range of studies and interventions to harmonise can strengthen causal understanding, rather than weaken it (Panter et al., 2019).

Panter et al. (2019) call on the principles of realist-review methods to develop a linear series of components that enable disparate evidence to be harmonised. These principles ask reviewers to consider these points:

- The function of an intervention: What an intervention is intended to do (for example, improve pedestrian safety).
- The context of an intervention: What social, physical, organisational or political characteristics exist where the intervention takes place (for example, geographical features of an intervention's location or social characteristics of its intended audience).
- The changes in process or reasoning that occur: How people or organisations respond to the intervention (for example, changes in people's perceptions of the safety or acceptability of walking as a mode of transport).
- The outcomes produced: What outcomes result from the intervention (for example, individuals or population groups changing their behaviour).

The realist-review method of synthesising evidence supports the mosaic approach (see section 1.2.3), because it allows for 'messy' complexity in evidence that is being harmonised. In these situations where there is no clear way forward, heuristics can help decision-making (Ivory et al., 2013). A heuristic is a 'rule of thumb' that focuses on practice rather than purely theory. It consists of a set of rules that aid decision-making in complex situations or when information is incomplete (Marewski & Gigerenzer, 2012; Sullivan, 2009). Rather than giving a precise, definitive answer, an accessible and understandable heuristic is designed to help people reach a 'good enough' answer (Marewski & Gigerenzer, 2012). It encourages users to take a step back from all the information, to consider what is most important rather than get bogged down in the detail (Marewski & Gigerenzer, 2012).

The realist-review components that Panter et al. (2019) use are a useful starting point for a heuristic that will help decision-makers harmonise disparate evidence.

2.3   Workshop with stakeholders

To help us develop the requirements for the CETI, we invited the catalogue's intended end users to take part in an online, interactive workshop. From the workshop, we wanted to find out why users would find a catalogue useful and what features would be most helpful to them. Twenty-four people attended the workshop, including representatives from NZTA, Te Manatū Waka Ministry of Transport and local authorities, and independent researchers.

During the workshop, we ran two small-group exercises using the online interactive Miro platform.[12] Participants were invited to post virtual sticky notes in response to prompts, and take part in small-group discussions.

Exercise 1: What makes a piece of evidence useable?

In their small groups, we prompted participants to consider what attributes and qualities should inform the standards for evidence that we include in the catalogue. We asked them to consider and discuss:

- types of evidence
- types of interventions

---

[12] Miro is an online interactive whiteboard. For more information, visit https://miro.com

- different audiences
- positive and negative attributes.

Exercise 2: What features, functions and qualities would make the catalogue work?

In their small groups, we prompted participants to consider what features would be desirable for a catalogue of evidence. We asked them to consider and discuss which features, functions and qualities would:

- make the catalogue easy to navigate
- give users confidence in the results
- provide users with an effective search experience
- make using the catalogue an enjoyable experience.

Group discussion

At the workshop, we shared the results of our review of existing catalogues, describing their key features and highlighting commonalities and differences between them.

We ended the workshop with a recap and discussion of the outcomes of Exercises 1 and 2.

2.3.1 Workshop results

We reviewed all the virtual sticky notes on the Miro board from Exercises 1 and 2 and grouped the content into thematic clusters.

Exercise 1: What makes a piece of evidence useable?

- **Searchable structure and language**: To enable evidence to be searchable, the catalogue needs to use consistent and comparable vocabulary and key information (such as dates).
- **Robustness and quality**: For users to trust the evidence, they need information about the robustness and quality of the research, evaluation methodology and inferences made from the results. Information about the source of reports is also important.
- **Clarity**: To enable a range of users to easily understand the evidence, the catalogue needs to limit technical jargon and ensure that methods and outcomes are well articulated.
- **Relevance**: To allow users to meaningfully compare studies and assess how applicable and relevant they are to a particular situation, evidence needs to include information about a study's context (such as location, sociodemographic and cultural factors).
- **Content**: Reflecting users' diverse information needs, the catalogue should cover a broad range of outcomes and interventions, if it is to be useful.

Exercise 2: What features, functions and qualities would make the catalogue work?

- Search mechanisms that cope with multiple disciplines and technical language. 'Fuzzy' searching capability that gives approximate or 'near enough' matches.
- Search outputs that can be shared, exported and represented visually.
- Visibility of current evidence gaps.
- Capability to summarise or synthesise evidence to some degree, so that users can make sense across evidence.
- Connections within the catalogue.

- Connections to wider evidence bases.

2.4   Summary of requirements

By reviewing existing catalogues (see section 2.1), reviewing best practice for cataloguing and managing evidence (see section 2.2) and holding a workshop with stakeholders (see section 2.3), we were able to answer the two questions we posed to understand what a 'good' catalogue would look like.

2.4.1   What information is needed to evaluate the effectiveness of interventions?

- The catalogue needs to be **comprehensive**. Evidence needs to include all transport outcomes and interventions (this includes policy, infrastructure and behavioural interventions), so that the catalogue covers the whole transport system.

- The catalogue needs to provide information on the **context** of interventions, so that users can make inferences about the results of interventions and understand their relevance to different transport settings. Mapping information, through spatial layers, is one way of adding context. It can show clusters of different types of interventions across countries and continents. A source map can show underlying urban form, key features (such as proximity to urban centres), transport infrastructure and community amenities.

- The catalogue needs to be **diverse**. It needs to include studies that use qualitative and quantitative methods, to provide the level of insight that users need and to enable them to understand causes. It needs to cover different types of assessment methods and reporting, to develop a 'mosaic' of evidence.

- The catalogue needs to relate to **New Zealand's decision-making process** around transport, to be useful for transport practitioners. It needs to link evidence to investment processes, such as benefits management guidance.[13]

- The catalogue needs to encourage users to compare evidence by the **intended function or purpose of interventions**, as this increases causal understanding across a mosaic of evidence.

- The catalogue needs to include **expert interpretation** to aid users. Search engines and reference libraries already exist to collate evidence. The added value of the catalogues we reviewed is that they extract and harmonise information and provide users with summaries and syntheses by topic experts.

- The catalogue needs to include **relevant information**, including intervention inputs (such as costs) and outputs (such as results). For evaluations of interventions' effectiveness and impact, the catalogue needs to include information about what goes into an intervention and what happened afterwards.

2.4.2   What form does the information need to be in?

- The catalogue needs an **organising hierarchy**. The structure needs to be defined through a taxonomy and content needs to be classified in ways that reflect the catalogue's purpose and intended audience.

- The evidence needs to be **defined**. The catalogue needs established standards about what counts as an intervention and as evidence.

- The catalogue needs to enable **comparisons**. Using standard formats and definitions enables evidence to be harmonised and compared with each other, and evidence of interventions to be compared with potential intervention sites.

---

[13] For more information, visit www.nzta.govt.nz/planning-and-investment/learning-and-resources/benefits-management-guidance

- The catalogue needs to be **documented**. The catalogue's standards for evidence and metadata need to be available and transparent.

- The catalogue needs to be **searchable**. The catalogue's form and structure need to facilitate searching, to answer questions from multiple perspectives. These will range from an overarching perspective of transport outcomes through to specific questions about how a reported intervention relates to others in the catalogue.

- The catalogue needs to show **gaps in evidence**. Understanding absent or inconsistent evidence is critical to improving the evidence base and, ultimately, our understanding of what works in transport. It is also important for users to see what is not yet known.

- The catalogue needs to use **consistent language**, and provide any necessary definitions, so that a range of users can use it.

- The catalogue needs to be **user friendly**, so that a range of users can easily use it to access evidence. Its user interface needs to adhere to useability principles, so it is intuitive.

3 Step 2: Designing and reviewing a prototype catalogue of evidence

The second step in the project was designing the CETI, and building and reviewing a digital prototype to demonstrate the design features and functionalities. This step did not involve establishing a fully functioning digital catalogue.

Step 2 involved:

- developing the catalogue framework and hierarchical structure, including evidence standards and metadata standards
- building a prototype catalogue, using a sample of sources to design the catalogue's categories and filters
- designing a digital interface
- developing a heuristic for harmonising results from diverse sources.

3.1 Catalogue design

3.1.1 Catalogue framework

The catalogue framework creates a hierarchical structure that is illustrated in Figure 3-1. The structure has five levels that are used to organise information:

Level 1: Historical window.

Level 2: Policy outcomes.

Level 3: Functions (changes to).

Level 4: Intervention types.

Level 5: Sources.

The catalogue framework responds to the requirements we identified (see section 2) in several ways.

3.1.1.1 An organising hierarchy that reflects decision-making processes

The top of the hierarchy is a **historical window** (level 1), which adds a temporal aspect to the taxonomy. This means that, as policies change, new transport policy outcomes can be added and linked to evidence in the catalogue, and information can be filtered by historical policy periods.

Level 2 of the structure is the **policy outcomes**, which align with the five **transport outcomes** in Te Manatū Waka Ministry of Transport's Outcomes Framework.[14]

The 12 **functions** in the structure (level 3) reflect the 12 benefit clusters in the Land Transport Benefits Framework.[15] The functions relate to the changes that interventions intend to make, such as changing the level of user safety in the transport system.

The functions contribute to change by using specific **intervention types** (such as low-speed zones) (level 4).

---

[14] For more information, visit www.transport.govt.nz/area-of-interest/strategy-and-direction/transport-outcomes-framework

[15] For more information, visit www.nzta.govt.nz/planning-and-investment/learning-and-resources/benefits-management-guidance/the-land-transport-benefits-framework/introduction-to-the-benefits-framework

At the bottom of the hierarchy are individual **sources** of evidence (level 5), which inform the choice of intervention type.

Each source of evidence (level 5) is categorised by the intervention type it relates to (level 4), the function it corresponds to (level 3), the policy outcome it contributes to (level 2) and the historical period that the policy outcome comes from (level 1).

Figure 3-1   Framework of the catalogue of evidence for transport interventions
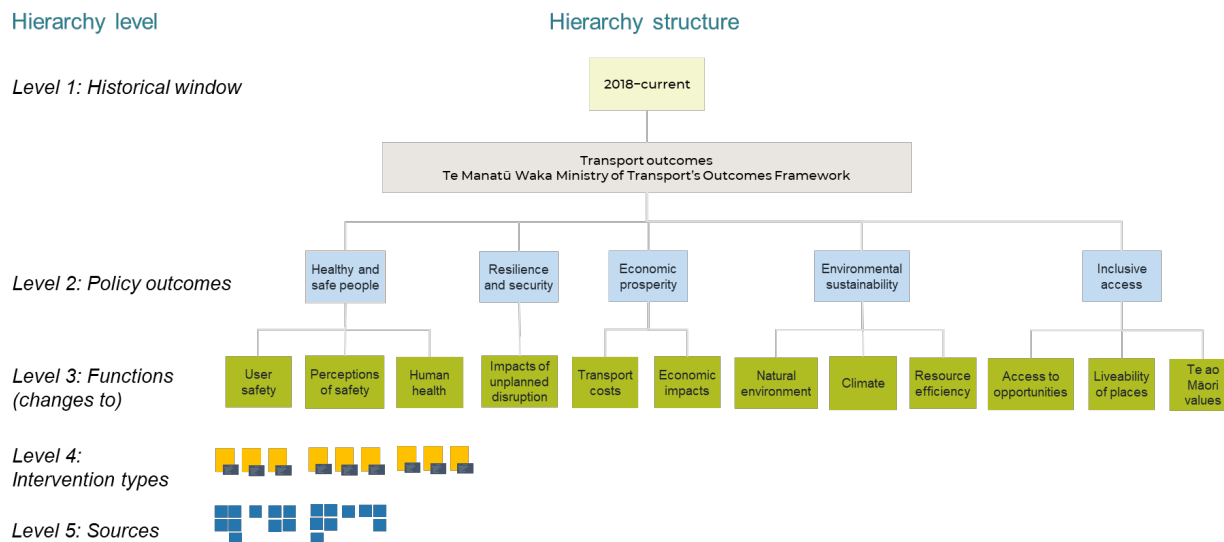
3.1.1.2   A catalogue that is searchable and enables comparisons

The sources of evidence in level 5 of the hierarchy are the building blocks of the catalogue. They include individual items that report a transport intervention, such as academic papers, research reports and project evaluations. The catalogue operates by categorising each source within each of the other four levels in the hierarchical structure, and using filters for search purposes. The categories and filters extract information from sources to answer users' queries. A single source can be linked to multiple intervention types, functions, policy outcomes or historical windows. For example, an evaluation of a mode-shift project could be linked to changes to the human health function and the changes to access to opportunities function. Conversely, a low-speed zone intervention type will be linked with multiple sources.

The catalogue framework includes categories and filters that the system uses to extract information from individual sources about a reported intervention (see Figure 3-2). The information they extract includes:
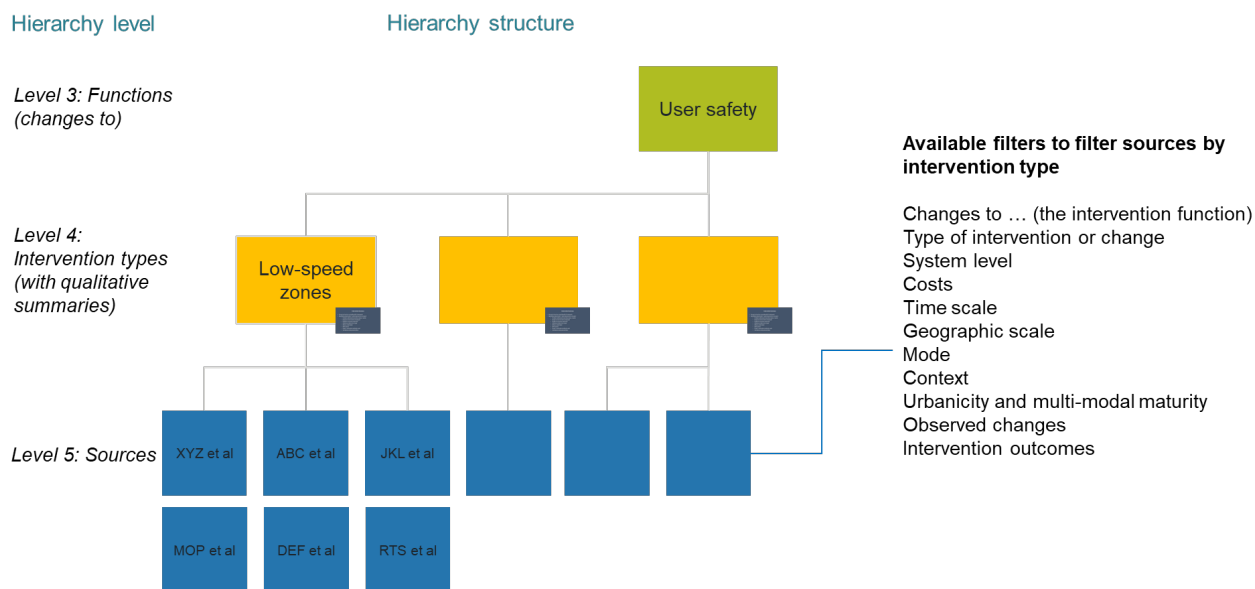
- the changes that were made
- the way the intervention was categorised (for example, infrastructure)
- the level of the transport system affected (for example, national or regional)
- the costs involved
- the time scale and follow-up period
- the geographical scale (for example, neighbourhood)
- the transport modes involved (such as walking or public transport)
- contextual factors (for example, the country, population groups or cultures)
- the level of urbanicity and transport maturity
- the changes and outcomes observed following the intervention.

Figure 3-3 illustrates the filters used to filter sources by their type. These include:

- citation details (such as, when and where a source was published, authors and abstract)
- the type of source (such as, journal article or peer-reviewed report)
- the type of data used to evaluate an intervention
- SMS rating[16]
- the assessment type (such as, randomised controlled trial or case study).

The full range of categories and filters is contained in Appendix A. When we developed the categories and filters, we sought a pragmatic and practical balance between having comprehensive coverage, but with sufficient detail. The framework, therefore, includes all transport outcomes, intervention types and evaluation types.

Figure 3-2   Filters available to filter sources by intervention type

Available filters to filter sources by intervention type

Changes to … (the intervention function)
Type of intervention or change
System level
Costs
Time scale
Geographic scale
Mode
Context
Urbanicity and multi-modal maturity
Observed changes
Intervention outcomes

---

[16] The catalogue uses the Maryland Scientific Methods (SMS) scale (What Works Centre for Local Economic Growth, 2015) to score each source for the extent to which it controls potential bias.

Figure 3-3   Filters available to filter sources by type
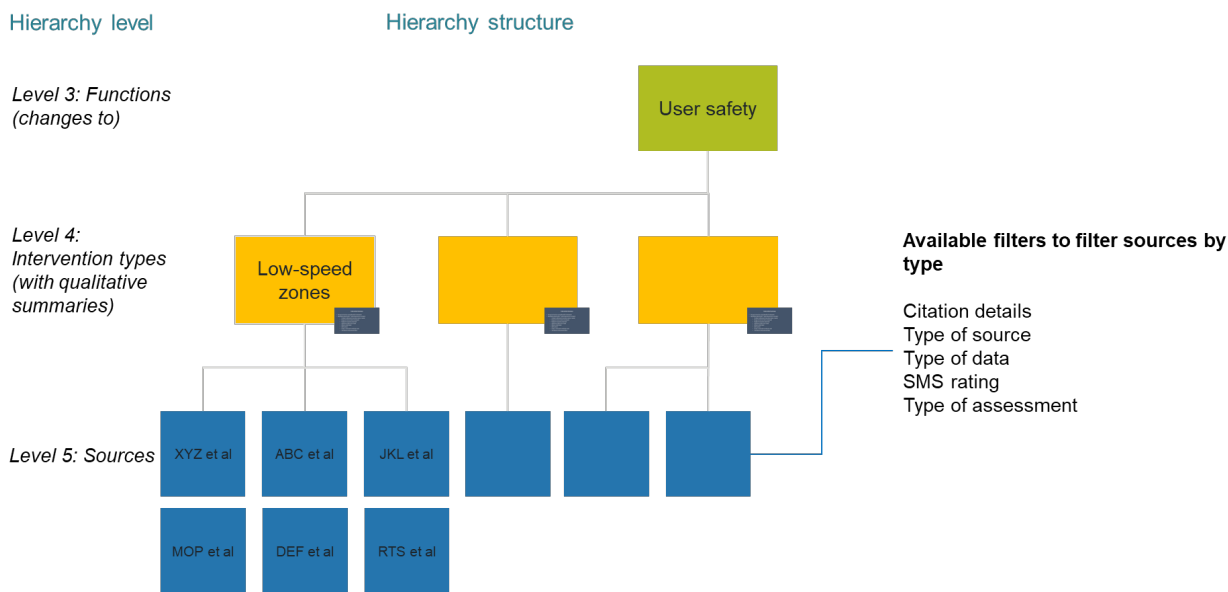
3.1.1.3   A catalogue that provides expert interpretation to aid users

Catalogues like SafetyCube synthesise quantitative studies to estimate the cost-effectiveness of interventions. However, this requires analysis and expertise that may not be feasible for NZTA's catalogue. This synthesis can also privilege certain types of evidence (such as quantitative evaluations of short-term, simple, single interventions), which mean users will be less inclined to consider a wider range of interventions. Therefore, our design does not depend on quantitative synthesis, although it enables the catalogue to capture as much information from sources as is practicable, so that synthesis is possible if it is required.

As the catalogue develops, a qualitative summary of each intervention type (such as low-speed zones) could be developed. These summaries are represented in Figure 3-2 and Figure 3-3 by small dark blue boxes next to the intervention types in level 4 of the hierarchy. The qualitative summaries could briefly describe:

- the problem addressed by the intervention type
- the range of interventions covered by the intervention type
- the range of reported outcomes for the intervention type
- inferences that can be made about the intervention type
- where the intervention type has been used
- who has used the intervention type
- the methods used to study or evaluate the intervention type
- the limitations and potential bias of evidence on the intervention type.

3.1.1.4   A catalogue that is user friendly

The way that sources are organised in a hierarchical structure gives users flexibility about how they perform searches. For example, our framework supports users to make a 'top-down' search that starts from a policy outcome (for example, 'What interventions contribute to improving safety?') and a 'bottom-up' search that locates a specific study within a group of intervention types or policy outcomes (for example, 'How does this piece of evidence add to the body of knowledge? Or 'If we did the same intervention, what outcomes and

functions could it contribute to?'). This flexibility allows users to explore specific policy targets (such as decarbonisation) by policy outcomes and functions.

3.1.2 Evidence standards

In this section, we explain the evidence standards that our design for a CETI uses. These standards are used to decide which evidence is included in the catalogue and how it is represented. The evidence standards are designed to be agnostic about the report type. Therefore, the catalogue includes traditional research reports published in academic journals, evaluation reports and operational reports, if they meet these evidence standards.

3.1.2.1 Interventions

For a source to be included in the catalogue, it needs to report on something that has changed, deliberately or unintentionally. This could include interventions that have led to specific changes to transport (such as introducing bus lanes) or interventions that have had an impact on transport (such as urban-regeneration programmes or disaster-related travel restrictions).

The catalogue includes micro-level interventions that are targeted at individuals, meso-level interventions that are targeted at groups or types of mobility (such as a type of journey purpose or a specific time of day when journeys occur) and population-level or macro-level interventions (such as broad environmental land-use changes) (Cleland et al., 2023). The catalogue excludes pure cross-sectional studies, if they cover only one time period. However, it includes a series of cross-sectional studies over the course of an intervention, if they are reported as an evaluation.

3.1.2.2 Assessment types

In keeping with the mosaic approach (Ogilvie et al., 2020), the catalogue includes interventions that are evaluated and reported on in different ways. It classifies sources as 'quantitative', 'qualitative' or 'mixed methods', which are the same classifications that SafetyCube uses. Each classification has a definition and examples of alternative names.

3.1.2.3 Sources

While the catalogue deliberately includes a range of evaluation types, to be included a source must describe the baseline (the situation before the intervention began or the change happened), what changed (the intervention or change) and what happened afterwards (the consequences of the change). Sources that describe only what happened after an intervention are not included as evidence of change.

3.1.2.4 Quality of evidence

While the catalogue includes all assessment types, it categorises them to indicate how robust their evaluation is. The catalogue uses the Maryland Scientific Methods (SMS) scale (What Works Centre for Local Economic Growth, 2015) to score each source for the extent to which it controls potential bias.

The SMS uses a five-point scale. A score of 1 is given to simple cross-sectional correlations or before-and-after comparisons without a control or untreated group; a score of 5 is given to randomised controlled trials. The scale is a relatively simple process to determine quality by reviewing the methods and results section of reports. However, assessors need research expertise and training to use it. If resources allow, NZTA could

consider using more comprehensive and rigorous methods to assess evaluation quality, such as the Quality Assessment Tool for Quantitative Studies,[17] which takes 30 to 60 minutes to assess each source.

3.1.2.5 Relevance criteria

To give users an indication of how relevant an intervention may be to New Zealand's urban settlements, we created a category called 'Relevance to NZ'. To complete this category, the person assessing the evidence needs to know about the New Zealand context, notably our urban settlement tiers[18] and the type of transport infrastructure that characterises the tiers. For example, a tier 1 urban settlement is likely to have a more comprehensive public-transport network than a tier 3 settlement. Tier 1 settlements are also more likely to have congestion issues than smaller settlements. We have added qualifiers to this category, recognising that, for example, a mode-shift intervention may be relevant only where public- and active-transport alternatives already exist.

The relevance criteria are necessarily crude, as detailed assessments are unlikely to be within the catalogue's scope, although they could be considered for selected topics. Instead, the criteria are designed to be a useful screening tool that users can follow with a more detailed assessment, if they are considering whether a specific intervention is relevant to their setting.

Additional relevance information, to assess whether an intervention would be valid in a different setting, can be provided through social-context variables (for example, sociodemographic information), the intervention setting (geographical and spatial factors) and the intervention type.

3.1.2.6 Impact criteria

The catalogue includes an overall assessment of the potential impact of an intervention. This combines information on the magnitude and direction (gains or losses) of the results reported in the source, with information on whether the change was intended to affect a targeted population group (such as travel plans that were intended to increase the number of school children walking to school) or the wider population (such as improvements to active-transport infrastructure that were intended to benefit all potential users).

3.1.2.7 Spatial location

The catalogue provides information on the spatial context of sources in three ways:

- It records country and settlement names, to enable free-text searching.
- It categorises sources by the geographical scale of the intervention (such as household, neighbourhood or region).
- It displays sources on a map, based on the coordinates of the settlement name provided in the source.

3.1.3 Metadata standards

The term 'metadata' means data that describe data (Hay, 2010) We developed metadata standards (see Table 3-1) to describe – from a scoping or planning perspective − which metadata are included in the catalogue and how they are organised to provide information about the evidence in the catalogue. The

---

[17] For more information, visit www.ephpp.ca/quality-assessment-tool-for-quantitative-studies

[18] The *National policy statement on urban development 2020* categorises New Zealand's urban environments into three tiers, based on population size and growth rate. Tier 1 includes Auckland, Tauranga, Hamilton, Wellington and Christchurch. Tier 2 includes Whangārei, Rotorua, New Plymouth, Napier, Hastings, Palmerston North, Nelson, Tasman, Queenstown and Dunedin. Tier 3 includes all other settlements (Ministry for the Environment, 2022).

metadata used in the CETI are based on the Hay (2010) Architecture Framework. In Table 3-1, we set out appropriate standards for metadata, based on what we have learnt from developing the prototype CETI. The table includes standards that NZTA will need to determine in the future, as it develops the catalogue.

Table 3-1   Metadata standards used in the catalogue of evidence for transport interventions

Area of interest		Standard
Data	Definitions and descriptions	Fields and categories should have plain-language names, definitions and descriptions, which are used consistently. When necessary, descriptions can be supplemented with technical language Descriptions and definitions must be sufficiently detailed that they support data-entry judgements and help users interpret the evidence
Data	Categories and free text	Categories must have exclusion criteria Free-text fields need character limits that are sufficiently big to include an abstract and executive summary
Activity	Managing users	Users' access to the catalogue needs to be defined and managed. This includes defining how the log-in process is established and overseen, how access levels to activities like data entry are determined, and how user behaviours are monitored and managed. Information about users (such as their email addresses) must be held securely Roles (such as managing users, entering data and searching) must be defined and managed
Activity	Selecting sources	The process to search for and select sources, to initially populate the catalogue and update it, needs to be agreed and specified. The process options include: • specifying which search terms to use in appropriate search engines for relevant disciplines that cover the range of transport outcomes • drawing from literature and evidence collections that NZTA or its stakeholders already use • extracting literature from published systematic reviews and meta-analyses • performing targeted literature searches as gaps emerge, using agreed search terms • asking the catalogue users for suggestions
Activity	Entering data	Processes for entering data into the catalogue must be consistent. A core set of fields can be 'required', so that an entry will not be saved until they are all completed. The core set can be determined as the catalogue develops Quality control: To ensure sources are entered into the catalogue consistently and thoroughly, data-entry personnel need to be trained, and their work − particularly fields requiring judgement − needs to be peer reviewed
Activity	Performing searches	Users can search all information in the catalogue using the search functions. The results of all searches can be viewed as a list or map The results of a search query can be viewed on the screen in a summary or expanded form. If the catalogue is further developed, results could be exported as a pdf or reported in a dashboard
Activity	Managing, evaluating and governing the catalogue	Responsibilities and processes for managing, evaluating and governing the catalogue need to be determined
Location	Online location	The catalogue will be hosted and managed by NZTA, and accessed through the NZTA website
Location	Geographic access	Access to the catalogue by an international audience needs to be decided

Area of interest		Standard
People	Administrators	User management: Responsibilities for managing users, and information about them, need to be defined and assigned, to ensure user information is secure and users have trust in the catalogue
		Data management: Responsibilities for managing data in the catalogue need to be defined and assigned, to ensure information quality and coverage is retained
People	Users	Users need to be defined. They could include NZTA staff, transport practitioners (for example, staff from ministries, local authorities and consultancies), non-transport professionals, advocacy organisations, community groups and members or the public. A wider group of users will need more support information
People	Stakeholders	Stakeholders need to be defined. They include people and organisations who will directly benefit from the catalogue (such as transport-investment decision-makers), people and organisations whose work contributes to the catalogue (evidence generators) and people and organisations who benefit from evidence-based decisions about the transport system
People	Governors	The governance arrangements for the catalogue need to be determined
Time	Evidence currency	Administrators need to decide if evidence is still current. Some evidence will remain relevant and current, while other findings may be overturned. The changing nature of transport may mean more recent evidence has greater currency. Time limits may be required, to restrict the catalogue to the most relevant and useful information
Time	Reviews and evaluations	The timing of reviews and evaluations needs to be defined, to maintain the quality of the data in the catalogue and users' interest in it
Motivation	Purpose and usefulness	The content of the catalogue needs to match users' and stakeholders' needs, to ensure it is used. Processes to engage with users and stakeholders, and evaluate and govern the catalogue need to be determined, to ensure the catalogue achieves its purpose and remains useful
Motivation	Updates	Requirements for updating material need to be established as the catalogue develops, to ensure it remains relevant for users. To ensure the catalogue's content is comprehensive, the content needs to be planned and periodically reviewed, to look for gaps. Updates should address gaps by searching for, and vetting, additional sources

Standards for how data are defined and described will need to be calibrated as the catalogue develops. For example, the CETI currently has nine categories of geographical scales. It may be appropriate to aggregate some (for example, combine individual and household), disaggregate others (for example, define more than one type of organisation) or add ones (such as, network).

Where possible, definitions and descriptions need to be relevant to New Zealand while recognising international terminology. For example, the CETI's social-context definitions will refer to the term 'ethnicity' but will include race-based terms as well, because these are used in some international literature.

Other useful categories are likely to emerge, as the catalogue is populated. They can be developed by analysing free-text entries and when new areas of content are added. As the catalogue develops, the optimal number of fields and categories will need to be agreed, to ensure the catalogue remains useful and useable.

Once the final categories are agreed, their definitions and descriptions will need to be further developed and tested with a range of users, to ensure they are clear.

3.1.4 Approach to harmonising evidence

The catalogue's design supports harmonising sources across the evidence base, so that the users can make reasonable inferences about an intervention's effectiveness. By extracting key information from sources, the

catalogue helps users compare similar interventions and make judgements about their methodological robustness.

Figure 3-4 illustrates how the degree of comparability or overlap between sources, or between a source and a new context for applying an intervention, relates to how much supplementary information and judgement users need to contribute to decide if an intervention is likely to be effective.

Figure 3-4   **Relationship between the comparability of sources and the amount of extra information needed to assess effectiveness**
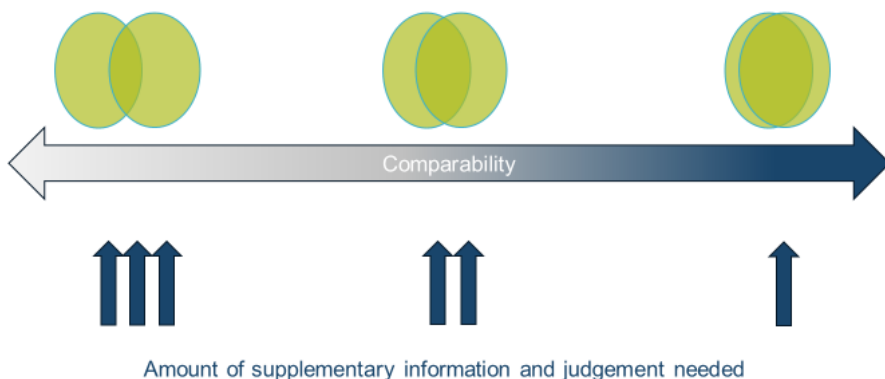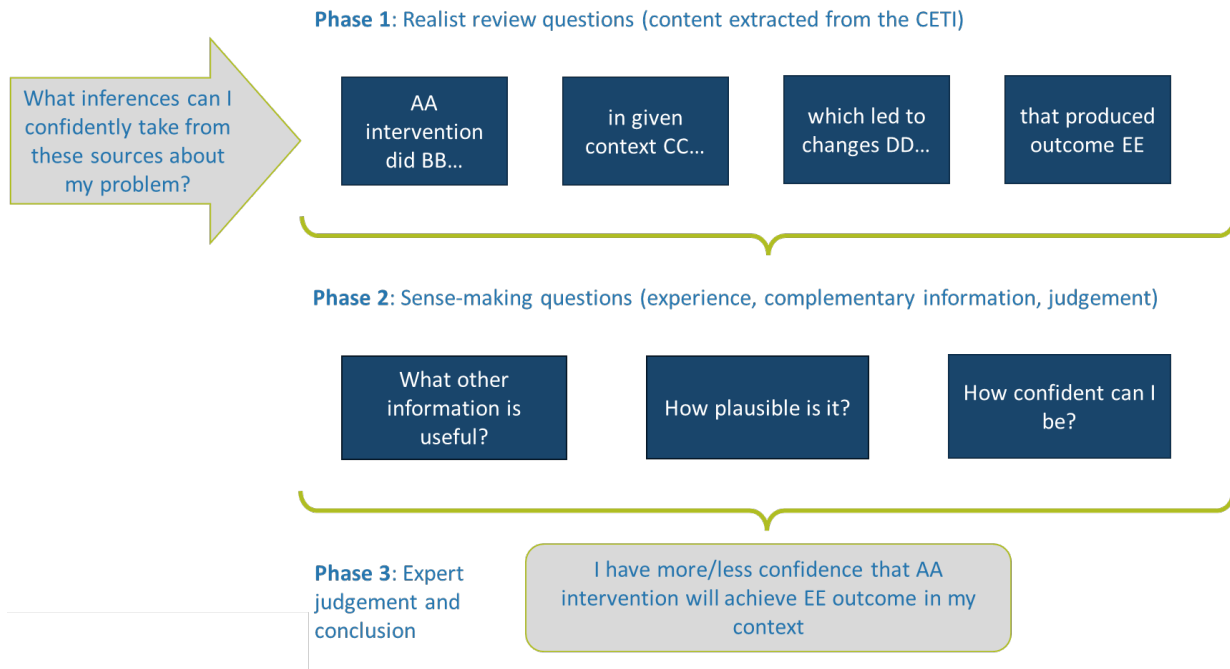

However, in our situation, it may often be impossible to closely compare sources, given the intended diversity and breadth of the catalogue, and stakeholders' desire to be able to translate reported findings to their specific context. Given this, we need to use a heuristic process that helps us compare sources, by seeking supplementary information and judgement from users.

The design of our CETI supports evidence to be harmonised in two ways:

- **Filters-based**: Standardising how information is extracted from non-identical sources creates a more cohesive database. It allows a user to compare sources by key factors related to specific interventions that have been evaluated and the context (reported and intended) of interventions, and to compare source-quality aspects such as whether they use the SMS scale, whether they were peer reviewed and what type of publication they were published in.

- **Heuristic-based**: A harmonising evidence heuristic gives users a process (see Figure 3-5) to ask the evidence base, and wider knowledge base, specific questions. When a user approaches diverse evidence, a question acts as a 'rule of thumb' that allows them to incorporate their expertise into the final judgement. The heuristic can be used to compare:

  – multiple sources within the CETI

  – sources in the CETI with external sources of evidence (such as a systematic review)

  – an evaluated intervention with a planned intervention.

Figure 3-5   The process used by the harmonising evidence heuristic in the CETI

Phase 1: Realist review questions (content extracted from the CETI)

> What inferences can I confidently take from these sources about my problem?

| AA intervention did BB… | in given context CC… | which led to changes DD… | that produced outcome EE |

Phase 2: Sense-making questions (experience, complementary information, judgement)

| What other information is useful? | How plausible is it? | How confident can I be? |

Phase 3: Expert judgement and conclusion

> I have more/less confidence that AA intervention will achieve EE outcome in my context

The harmonising evidence heuristic starts with the question: 'What inferences can I confidently take from these sources about my problem?' The heuristic then guides a user through three phases:

- **Phase 1: Realist-review questions**. In this phase, the user responds to questions that enable the system to extract evidence from the catalogue using filters.

- **Phase 2: Sense-making questions**. In this phase, the system asks the user to consider additional sources of information that may help them make sense of evidence extracted from the catalogue.

- **Phase 3: Expert judgement and conclusion**. In the final phase, the user applies their expert judgement to the results of Phase 1 and Phase 2 to decide how confident they are that an intervention will produce the desired changes in their context.

Table 3-2 lists the realist-review questions the heuristic uses in Phase 1 and the sense-making questions it uses in Phase 2, and gives details of the information it uses to answer the questions.

Table 3-2   Questions used by the harmonising evidence heuristic in the CETI

Phase 1: Realist-review questions	
Heuristic question	**Filter**
AA intervention did BB…	Primary and/or secondary change type (categories) AND mechanism (type & description)
in given context CC…	Social context factors (demographics etc), urbanicity, geographical scale. Includes factors that can be gleaned from the spatial location but are not included in the source details (such as climate zones, population density, topographical features and sociopolitical factors). This filter would require spatial layers to be included in the map
which led to changes DD…	Change observed (category)
that produced outcome EE (and other outcomes)	Transport outcomes (category) Intervention function (category)

Phase 2: Sense-making questions	
Heuristic question	Examples of information
What other information is useful?	Relevant previous experiences, technical best practice, novel applications from other disciplines, relevant case studies, cross-sectional studies, systematic reviews and meta-analyses
How plausible is it?	User judgement about how well it fits with theory or mechanisms; whether it makes sense with real-world observations, or why it is different; and how relevant it is to their issue and setting
How confident can I be?	Publication source, peer reviews, SMS scale result for bias control and user judgement about the similarity to their context and problem
Phase 3: Expert judgement and conclusion	
I have more/less confidence that AA intervention will achieve outcome EE (and other outcomes) in my context	

The harmonising evidence heuristic combines information extracted from the catalogue with other knowledge sources in a semi-structured way. It allows users to consider a range of information to assess whether a study will be valid in their context, rather than rely purely on published evidence. For example, a user could consider whether New Zealand has the required regulatory framework or resource capacity to implement a particular intervention. Ultimately, the heuristic calls on a user's judgement, but it aims to increase the transparency and quality of that judgement, thereby increasing the user's confidence in the inferences they make about an intervention.

3.2   Catalogue prototype

We built a live digital prototype that meets the requirements described in section 3.1. We had intended to use one of WSP's existing digital-information platforms; however, the complexity and comprehensiveness of the CETI framework and requirements would have made this platform too onerous for users. Instead, we developed a bespoke platform to demonstrate the framework's functions and trial as many of the desired features as possible, within the project budget and timelines.

We designed the prototype to meet user-interface useability standards and best practice for cataloguing and using evidence, to make the catalogue easy for users to use. We populated it with a small sample of sources that were manually created from two previous literature reviews (Pacheco & Ivory, 2023; Thomas et al., 2022). We used a spreadsheet to transform the framework into categories, filters and values, and their associated definitions and descriptions (see Appendix A, Definitions). We populated the spreadsheet with information we extracted from the sources, and judgements based on our specialist knowledge. We then used the structure and content of Appendix A to develop the digital prototype (the assessments and judgements in Appendix A are intended solely to develop the prototype and demonstrate its capability). Once we had built the prototype, we manually entered data from the spreadsheet (see Appendix A, Sources) into the system, for demonstration purposes.
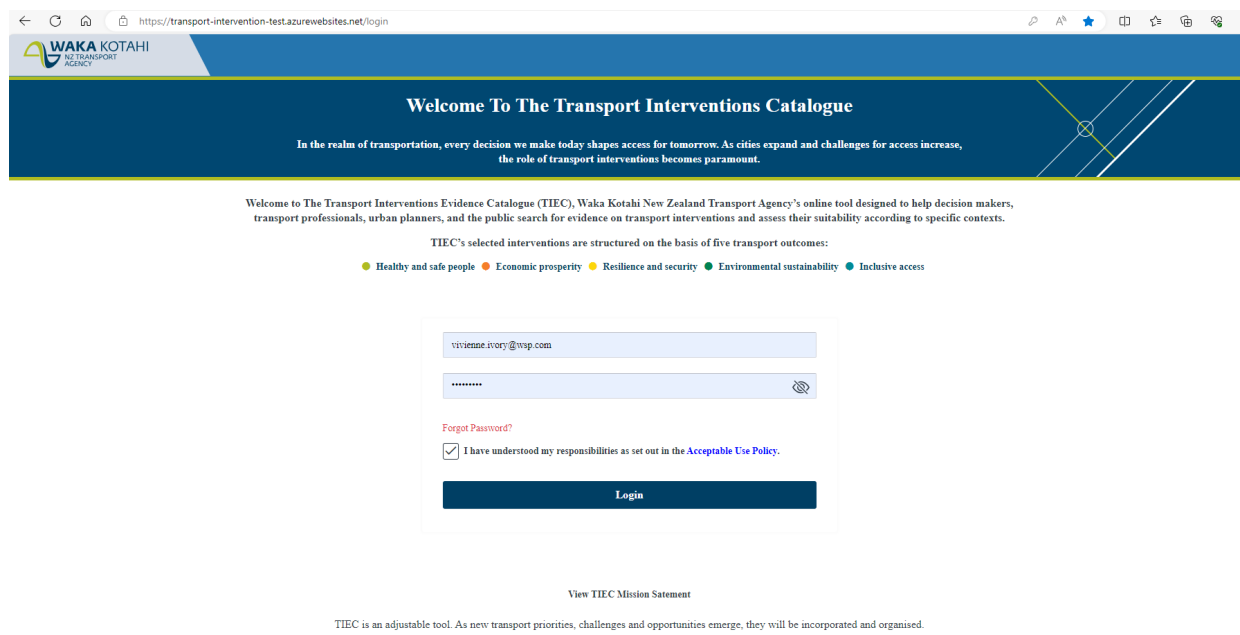
See Appendix B for further details of the design and metadata structure.

In the following sections, we describe the prototype CETI's features and functions.

3.2.1   Welcome page

The welcome page (see Figure 3-7) provides high-level information about the catalogue. It can be configured so that users need to log in.

Figure 3-6   Screenshot of the CETI welcome page

Note: Since the prototype catalogue was developed, the catalogue's title has been changed to a catalogue of evidence for transport interventions.

3.2.2   User management page

The extent to which users can access sources is known as 'managed access'. Managed access can be defined in the catalogue, so that administrators can give users approval to view or search sources, enter data, or activate and manage other users.

Figure 3-8 is a screenshot of the CETI Users page. It demonstrates how user information can be displayed and ordered alphabetically by name, email address, category (such as administrator, NZTA user or external user) and status (active or non-active). On the right-hand side, the administrator can choose user-management actions, such as editing user information, or archiving or deleting users.

Figure 3-9 is a screenshot of an individual CETI user's page. It shows how the administrator can assign roles to a user (such as administrator or standard user) and monitor a user's activity.

Figure 3-7   Screenshot of the CETI Users page
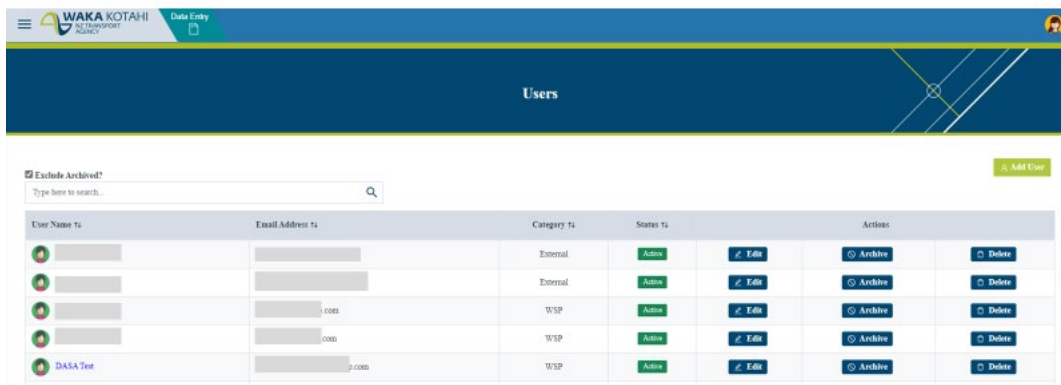
Figure 3-8   Screenshot of an individual CETI user's page
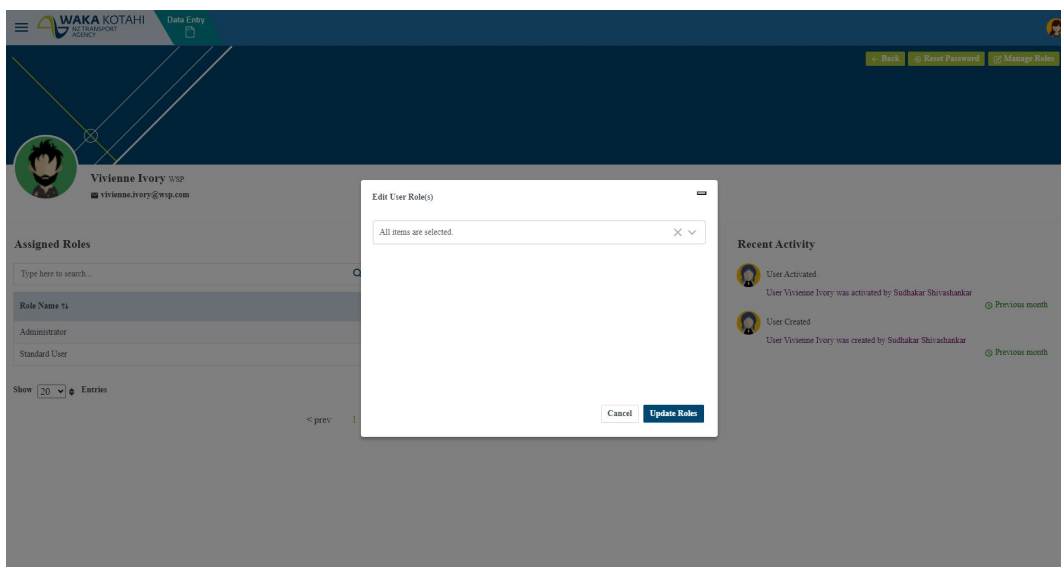

3.2.3   Data-entry page

The page for entering source data is split into four sections: Source Overview, Intervention Details, Evaluation Quality and Transport Outcomes (see Figure 3-10). The fields in these sections are populated with data from dropdown categories (for example, the Adjusted Maryland Scientific Methods Scale field), selection buttons (for example, the Type of Data field) and free text (for example, the Assessment Description field). An overall rating for Impact or Quality can be given by clicking on the number of stars.

A source can be assigned to multiple categories for one field or restricted to a single category. For example, a source can be relevant to more than one of the Transport Outcomes categories, but only have one geographical scale.

Fields can be set as optional or required. If a field is required, the entry for a source will not be saved until the field has been completed.

As the catalogue is populated with more sources, users will be assisted with the free-text fields through an autocomplete function.

Figure 3-9   Screenshot of the source data-entry page
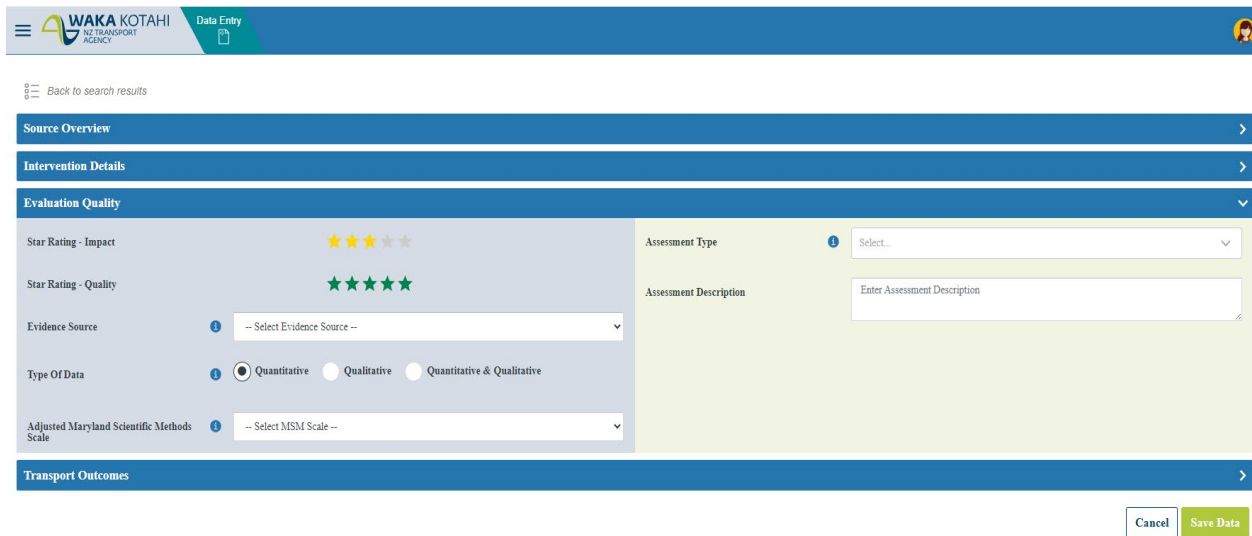
Information is extracted from a source in several ways, which require different levels of effort and skill (see Table 3-3). Some fields (such as Authors) can be populated easily, by taking data directly from the source. The data to populate other fields (such as Location) may need to be searched for in the source document. Completing other fields (such as Transport Outcomes) will need interpretation, assessment and judgement.

Some fields have an information button that provides a definition and description, which is based on the metadata (see Appendix B, Definitions). It is possible to link the information button with a support tool, such as the classification for urban settlement tiers[19] or the SMS manual (What Works Centre for Local Economic Growth, 2015).

Table 3-3   Methods of extracting data from sources

Extraction method	Data field example	Effort and expertise required
Copy and paste	Title Authors	Low
Find	Location Costs Included	
Interpret	Geographical Scale Assessment Type	Medium
Assess	SMS Transport Outcomes	
Judgement	Relevance to NZ settlement tiers	High

---

[19] The *National policy statement on urban development 2020* categorises New Zealand's urban environments into three tiers, based on population size and growth rate. Tier 1 includes Auckland, Tauranga, Hamilton, Wellington and Christchurch. Tier 2 includes Whangārei, Rotorua, New Plymouth, Napier, Hastings, Palmerston North, Nelson, Tasman, Queenstown and Dunedin. Tier 3 includes all other settlements (Ministry for the Environment, 2022).

3.2.4   Search features

Users can search sources in a list or map view, and can toggle between the two views.

3.2.4.1   List view

In list view, sources are displayed in a list, with several columns of data fields (the data fields can be determined later) (see Figure 3-11). The order in which items are displayed can be toggled by the column heading (for example, ordering alphabetically by Source Title name or categorically by Resource Type). Once a user enters a search term, the list is updated to show items that match the search result.

Figure 3-10   Screenshot of the list view in the CETI

Each row in the list can be expanded to display a summary view (see Figure 3-12). The summary view presents further information about the source in four tabs:

- Source Overview.
- Intervention Details.
- Evaluation Quality.
- Transport Outcomes.

Users have the option of expanding the summary view to a full view, which shows all the information available on the source.

The content of the summary view could be refined to include components of the harmonising evidence heuristic.

Figure 3-11   Screenshot of the summary view (list view) in the CETI
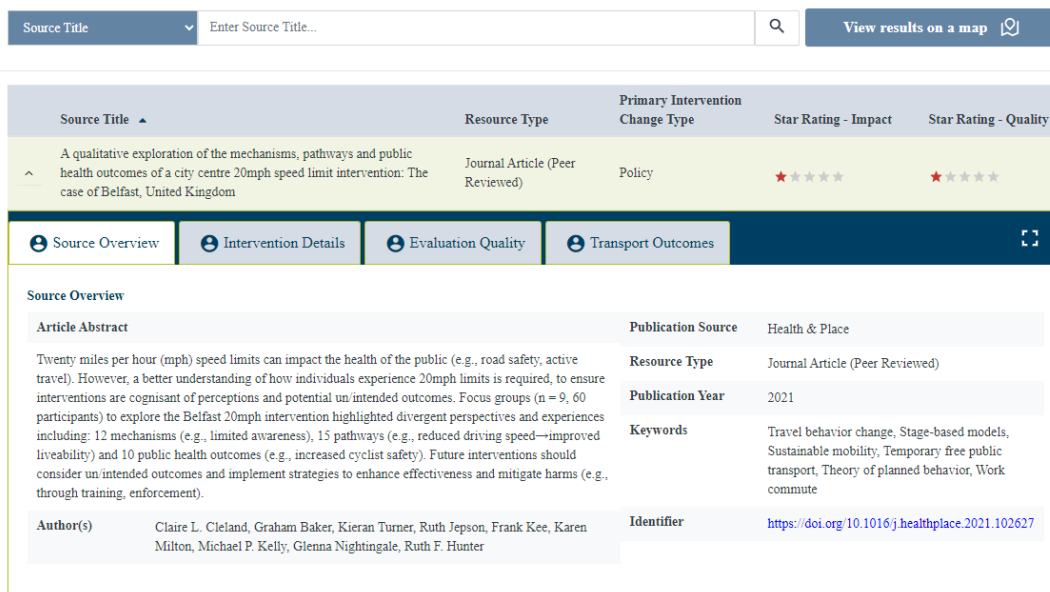
3.2.4.2   Map view

The map view shows where sources are located, based on the coordinates entered for each source (see Figure 3-13). Users can use the map to add contextual information about an intervention through spatial layers (such as climate zone, population density and walkability score), if they are available.

When a user conducts a search, the results of the search are shown on the map as pins (one pin for each source). Clicking on a pin shows the details of that source.

Users can toggle between the list and map views without losing the search results.

Figure 3-12   Screenshot of the map view in the CETI
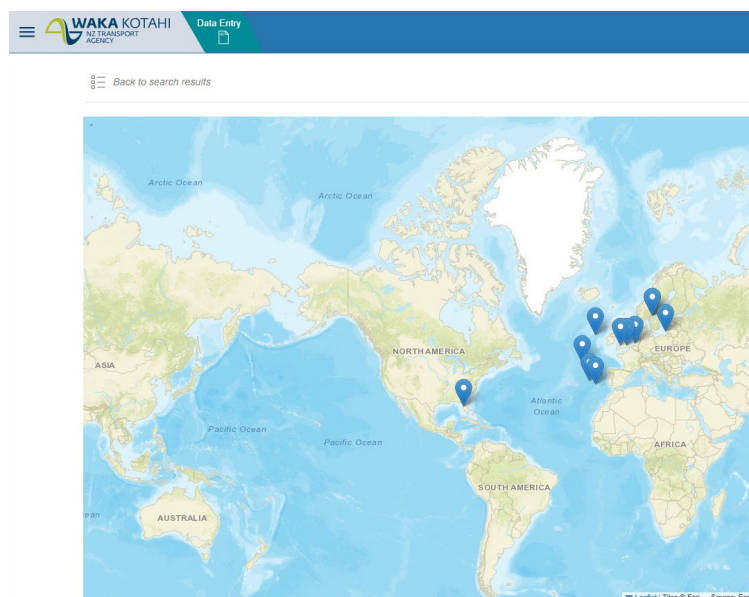
3.2.4.3   Search function

The search function consists of two parts:

1. Selecting a field (such as Assessment Type) from a dropdown box (see Figure 3-14).

2. Selecting a category (such as Before and after study) from a dropdown box (see Figure 3-15).

The user defines the order in which a search occurs, which means that their query can start anywhere within the CETI hierarchy. For example, they could start a search by selecting the 'Inclusive access' category in the Transport Outcomes field and use free text to search in the Location field. Or they could start a search by entering free text in the Location field, and then selecting a category in the Transport Outcomes field. Searching more than one field is called a 'consecutive search' (see Figure 3-16). Users can also add additional categories to a search (such as two categories in the Transport Outcomes field) (see Figure 3-16).

In the prototype, we have set searches as 'additive', meaning x AND y = result. A user can remove a search query and result at any time, by clicking on the 'x'. NZTA can consider whether a 'combining' search, meaning x OR Y= result, is more useful for users, or whether to enable both search types.

A category search can be 'exclusive' or 'multi-select', depending on how categories are specified when they are entered into the catalogue. For example, a user can select multiple Transport Outcomes categories (a multi-select search) but only one SMS rating (an exclusive search).

Figure 3-13   Screenshot of the search function in the CETI: Selecting a field
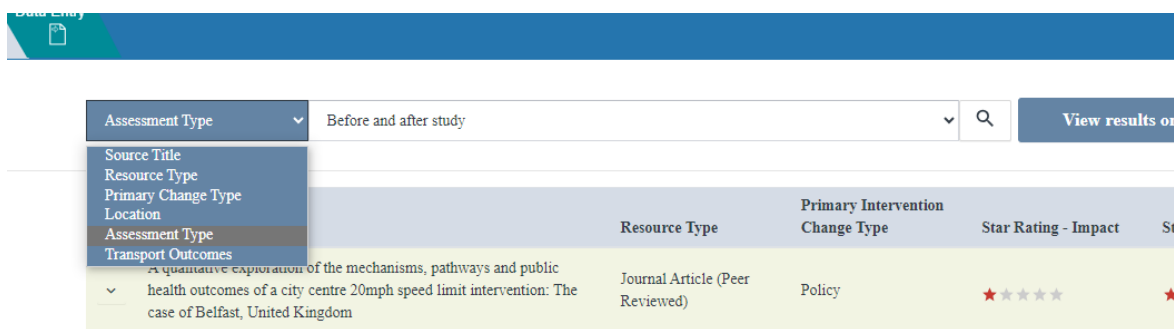
Figure 3-14   Screenshot of the search function in the CETI: Selecting a category
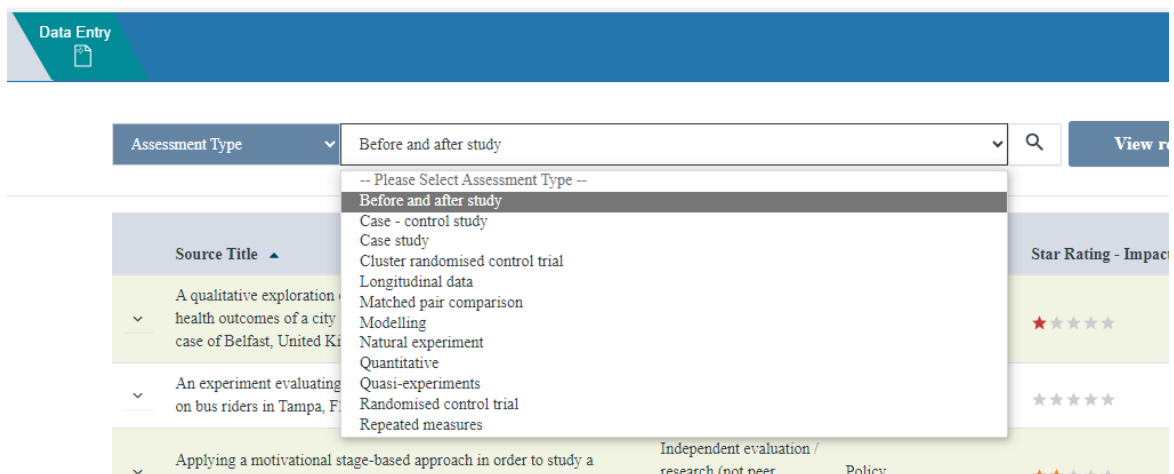
Figure 3-15   Screenshot of the search function in the CETI: Performing a consecutive search

We have designed the search interface so that these additional user-focused features can be incorporated into it, as required:

- The option for a user to select sources they want to save from a search.

- The option for a user to download search results as a shareable pdf.

- The option for a user to save their search results to a dashboard summary that shows, for example, the number of intervention types by Mode and Transport Outcomes fields.

3.3   User review

3.3.1   User-review method

To review the prototype, we interviewed four potential users of the CETI: three NZTA staff and one council officer. During the interviews, we demonstrated the CETI's features and prompted participants to reflect on how useful these features are, what limitations they have and what changes we could make to the catalogue in future.

3.3.2   User-review results

Overall, our participants see the catalogue is useful for these types of activities:

- responding to Official Information Act 1982 requests

- providing councils with evidence to help them make decisions and design interventions

- supporting business cases with relevant evidence

- staying informed about what their peers are doing (especially through case studies)

- systematically capturing evidence for transport interventions across their organisation

- incorporating relevant research into wider evidence about transport interventions.

Table 3-4 gives a summary of the participants' observations and suggestions.

Table 3-4   Results of a review of the prototype CETI

Theme	User observation	User suggestion
Ease of use	**Searching**: After observing the CETI in action, participants found it easy and intuitive to search. The hierarchical structure means that users can select relevant fields and categories according to their interests	Provide initial guidance, through additional text or a demonstration video

| Theme	User observation	User suggestion
	Data entry: The process to enter data made sense to the participants, but they thought the large number of categories may be off putting. They appreciated the CETI's ability to set some 'required' data fields. Most fields will need an information pop-out, to help ensure sources are categorised consistently	Provide information about categories, to ensure sources and assessments are categorised consistently
	User management: Participants considered the ability to manage users' access, track their usage and, potentially, engage with them is valuable	Automate user approval as much as possible [a]
Comprehensiveness	All participants recognised how comprehensive the catalogue design is, and that it seeks to cover all transport outcomes, types of assessments (case studies as well as quantifiable studies), and domestic and international sources	See note [a]
Consistency	All participants talked about the need for data to be entered and assessed consistently, for users to have trust in the catalogue	Provide clear guidance on how to enter data and conduct searches
Effort required	Participants recognised it will take some effort to develop and maintain the catalogue. They were particularly concerned about the data-entry burden associated with developing and maintaining such a comprehensive catalogue	Consider which information is most valuable to users [a] Consider which data fields should be 'required' and which could be discretionary and still enable the catalogue to harmonise [a]
Assessment and judgement	Staff who enter data will need varying levels of expertise to make consistent, robust assessments and judgements	Train staff who enter data, and monitor the quality of data, to ensure data are entered accurately and consistently
Access and data entry	All participants thought it was valuable to offer inclusive access to the search function. They discussed giving users permission to enter data, but said this would need to be carefully moderated, to ensure that data are high quality and the range of sources entered remains broad and does not introduce bias into the catalogue	Govern how users can access the catalogue and enter data into it
Gaining and keeping council interest	All participants thought the catalogue will be of great interest and value to councils and other users. However, they said that gaining and maintaining their interest would depend on the initial quality and volume of sources in the catalogue and keeping the catalogue updated and relevant	Invest resources into developing the catalogue to a useful size before releasing it, and into maintaining it through ongoing data entry [a]
Mapping studies	If additional spatial layers can be added to the catalogue, participants thought that mapping sources would be useful. However, as maps are not intuitive to everyone, they saw value in keeping both map and list view options	Consider what spatial layers could be added, even if they are country specific
Star rating	Participants saw the simplistic, yet highly visual, star ratings as a potential risk, unless they are carefully framed. Attempting to include multiple parameters into one rating could introduce a bias when users use the ratings to compare studies	Consider whether a star-rating system is appropriate [a]

Note

[a] Participants agreed that a comprehensive catalogue would be very useful; however, they identified several areas that may need to be traded off. These are also discussed in section 4.

4 Step 3: Agreeing the scope of the catalogue of evidence with the steering group

The third, and final, step in the project was to present the catalogue design and prototype and results of the user review to the steering group and agree on the appropriate scope for the CETI, given the need to make trade-offs, based on feedback from potential users.

We held a workshop for the steering group where we demonstrated the prototype CETI and explained how it had been developed. Our user review had identified the trade-offs between achieving a comprehensive catalogue and expending the resource and effort that a comprehensive catalogue would require. These trade-offs approximately result in four catalogue options (see Figure 4-1).

Figure 4-1   Catalogue coverage trade-off options
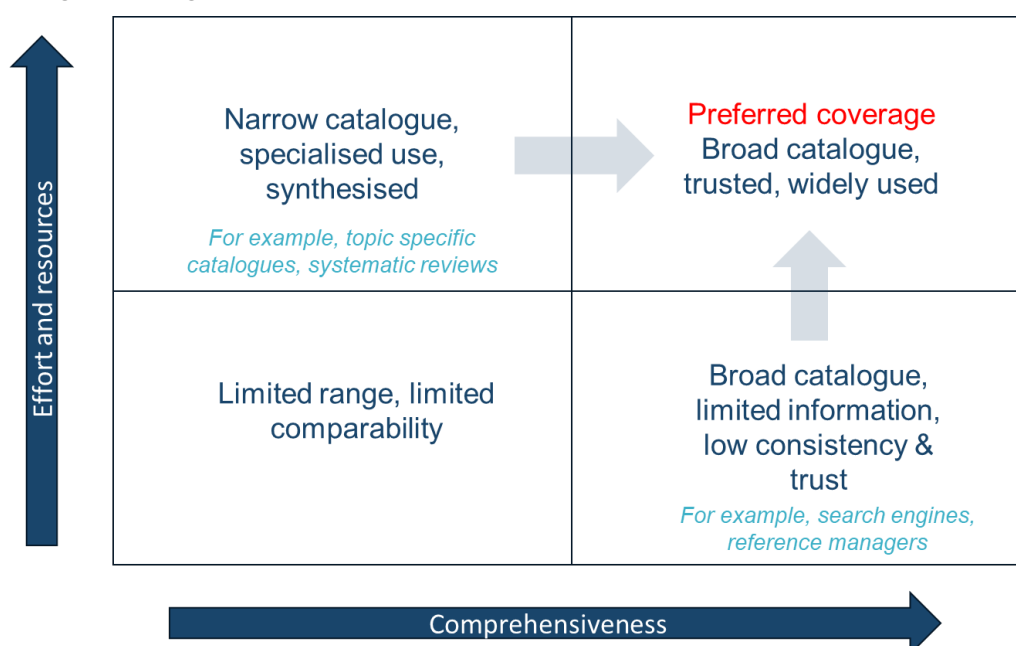
To weigh up the trade-offs, the steering group considered two main factors:

- What is necessary, in terms of
  - the core content that will add value to what is already available
  - the type of harmonisation and synthesis we need
  - the assessments and judgements we want
  - the range of transport outcomes, transport modes and assessment types we want to cover?
- What is realistic, given
  - the funding we have
  - the capacity we have
  - the capability we have?

The steering group's overall conclusion is that the CETI needs to fit into the top right-hand quadrant of Figure 4-1. However, the steering group also consider that existing tools that fit into the other quadrants could complement the CETI.

For example, in the bottom right-hand quadrant, search engines and reference managers (such as PubMed and Zotero) have relatively limited information and no assessments, but they do provide extensive coverage.

In the top left-hand quadrant, narrowly focused tools (such as SafetyCube, KonSULT and TCAD) are less comprehensive, but they do provide synthesis.

There is also an extensive, growing body of specific systematic reviews and meta-analyses in the transport sector that can complement the CETI's goal to provide a broad catalogue of evidence.

5   Recommendations

We have developed 10 recommendations for further developing the CETI, based on:

- our observations while we developed the prototype
- our insights from the user review
- our steering group's insights at the final workshop.

Our recommendations seek to balance the original scope we were given by NZTA with the requirements for the catalogue that we established in Step 1 (see section 2), the lessons we learnt from designing and developing the prototype (see section 3), the review by a sample of users (see section 3.3) and the insights of our steering group (see section 4).

Keep the coverage of the CETI sufficiently broad, so that harmonisation and additional synthesis and analysis are possible

The design for the CETI covers a broad range of transport outcomes, transport modes and assessment types, to align with the original project scope and the needs of stakeholders, and to enable data to be harmonised. While we were designing the catalogue, it became clear that the CETI could not, or should not, accommodate all these requirements. In particular, requiring the catalogue to assess cost-effectiveness and harmonise quantitative analyses of effectiveness could limit the diversity of evidence that the catalogue includes. However, it is possible −and necessary − to ensure that the CETI captures information in a way that allows these more detailed analyses to be undertaken in the future.

Further develop and test the language in the CETI, to ensure it is clear and used consistently

In the CETI design, we have used neutral terms, rather than discipline-specific terms. For example, we have used the terms 'intervention', 'instrument' and 'measure', which are more commonly described in sources as 'changes'. We have used the word 'source' to mean a piece of evidence, rather than standard research terms such as 'study', 'report' or 'publication'. However, the risk of using neutral language is that it introduces a new term that users may not understand. The CETI's category labels should be tested with non-technical users and users from different technical areas, to determine the most appropriate and easily understood terms.

Invest effort and resources into quickly getting the CETI to a useful size

It is important that the CETI reaches a useful size quickly, to ensure that councils are interested in and use it. Given the effort and resource required to enter data into the catalogue, an initial investment will be necessary to achieve this. This investment should extend to ensuring the CETI covers a broad range of transport outcomes, transport modes and assessment types, as this will reduce bias towards some types of studies, transport outcomes or interventions. With quality control, it is likely that artificial-intelligence tools could assess some of the sources, to speed up data entry.

Ensure that data are entered consistently into the CETI, so users are confident in the content

To make it easier to populate the catalogue, and to engage with its users, consider enabling two groups to enter data:

- **External users** can upload source details, such as citation details.

- **Trained administrators** can review and assess sources and upload all source data.

Administrators will need training and support materials, so that their data entries and assessments are consistent and valid. Given the wide range of source types, disciplines and study designs that will be represented in the CETI, they will need a high level of expertise to assess the quality of sources. They must also be familiar with the transport sector to assess a source and determine its classifications for, for example, the Transport Outcomes and Intervention Function fields.

Work towards the evidence in the CETI adequately covering all transport outcomes

It is critical that users can access evidence that is relevant to their area of interest. However, achieving a catalogue that covers a broad range of topics needs to be balanced with maintaining high-quality evidence. If the catalogue has a large proportion of sources that have a high potential for bias (for example, uncontrolled before-and-after studies) it may skew the evidence base, which would reduce users' confidence in it. On the other hand, if the CETI relies on a small number of sources with robust study designs, it could result in critical knowledge gaps. This would also frustrate users and may hinder investors. To ensure the CETI continues to meet users' needs, it will need a process to identify and address gaps in the substantive content and evidence quality.

Provide CETI users with access to full-text sources

Where possible, source entries should include a link to a source's full text with a DOI identifier. This may not always be possible, as some sources will be behind a paywall or have an outdated URL. NZTA should consider securing access to sources, so they can be stored or accessed by CETI subscribers. If this is not possible under licensing, a source entry should include sufficient information to enable users to make an adequate judgement about whether it is relevant to their business case or intervention design. In some cases, when a user needs more detailed information about a source, their organisation will need to buy the individual source or, in the case of an outdated URL, undertake further searching for it.

Include links in the CETI to complementary resources, such as search engines, systematic reviews and meta-analyses

The CETI should provide users with links to other knowledge repositories, such as specialist search engines, and the ever-growing body of evidence from high-quality systematic reviews and meta-analyses (full-text versions, where available). These resources often have a narrow focus, but their robust methodology means they can address users' questions about the effectiveness and transferability of an intervention. They also provide a useful entry point into published literature and may, therefore, be valuable for the initial population of the catalogue. The intervention summary (see 3.1.1.3) should include specific links and have a separate searchable page that covers systematic reviews and meta-analyses.

Focus the CETI content on what is most relevant for users

The CETI is designed to help users better understand if an intervention can work in New Zealand. Councils are most interested in what their peers are doing in New Zealand or similar jurisdictions. However, it could still be relevant to include sources about interventions from more diverse settings, as they may contribute to a deeper understanding of *how* to design an intervention to have the most impact. If the catalogue provides users with the means to usefully compare sources, international examples could also expand our knowledge of what is possible in New Zealand.

The design has not set any criteria for intervention context, but NZTA could consider screening sources to find content that is most relevant to New Zealand's transport context. For example, if an intervention's setting is a mega-city or a developing country, it will be less readily translatable to New Zealand.

Encourage new and better-quality evidence by setting standards for evidence included in the CETI

The CETI could encourage new evidence to be generated. First, it could make evidence gaps more visible. If the CETI includes a search-results dashboard, this could be used to identify gaps in the evidence base. While the CETI is designed to include operational evaluations, the necessary data for these evaluations are not always collected and reported, which restricts evaluations being used as valuable, local evidence. The CETI could set standards for NZTA research reports, operational reports and case-study evaluations, and NZTA could make meeting these standards a project requirement.

Avoid using overall ratings in the CETI, unless they will be easily understood

When we were defining the catalogue's requirements, some of the CETI's stakeholders suggested using a star-rating method. However, to get the rating method 'right' would require considerable testing, as the risk of a rating being misinterpreted is high. This risk is compounded because the rating gives a highly visible summary of the evidence, although this visibility is − of course − the purpose of a star rating.

6   Conclusions

Transport practitioners in New Zealand require better access to evidence to support their decision-making. They want to know more about how effective different kinds of interventions are, especially how effective an intervention could be at addressing their problem and how relevant it could be in their setting.

This report describes how we designed and prototyped a CETI that has structured and searchable information about transport interventions and their outcomes, to give New Zealand transport practitioners better access to evidence. To achieve this goal, we grounded the project in New Zealand's transport investment framework and used a 'mosaic' approach to provide the diversity of evidence that practitioners and NZTA need to answer their questions.

In our project, diverse evidence means that the design and prototype cover all transport outcomes, intervention types and evidence types. The CETI's diverse evidence base enables practitioners to understand what causes interventions to be effective and transferable to their setting, because they can extract data about an intervention's context and its specific outcomes. The design also standardises information from non-identical sources of evidence and uses a harmonising evidence heuristic to help users make inferences about findings and judge the degree of confidence they can place on the results.

The goal for the CETI is a catalogue of evidence on the effectiveness, cost, transferability and applicability of transport interventions. However, the desire to have a broad and comprehensive catalogue, and the nature of much of the evidence for transport interventions poses challenges. First, information types (such as information about costs, context and transport outcomes) are often missing or recorded inconsistently in sources. Second, methods that other catalogues use to systematically synthesise evidence are resource intensive and restrict the type of evidence a catalogue can include, which goes against our requirements.

Rather than restrict the CETI to a narrow range of evidence, our design ensures that, where possible, the catalogue gathers consistent information from sources in standardised ways, so that it can compare diverse evidence. This approach will, in future, mean the catalogue can offer more specialist synthesis using established methods (such as the cost-effectiveness analyses in SafetyCube).

We have designed the framework, evidence standards and metadata standards to ensure the catalogue:

- includes sources that cover all transport outcomes and interventions
- has standards for what and how evidence is included, so that evidence is machine searchable
- provides information on the rigour and robustness of the sources it contains
- clearly defines and describes its content
- adds value to its sources, by assessing their relevance to New Zealand urban settings.

The design incorporates the features, functions and qualities that best address user requirements. For example, using both free-text and categorical fields for entering data and searching means the CETI is designed to cope with multiple disciplines and technical language. To make evidence gaps more visible, the search outputs are designed to be downloaded or reported by a dashboard, although the prototype does not include this capability.

Despite how critical an intervention's cost and context are to evaluating its effectiveness, many sources provide limited information on these matters. Including these categories in our design highlights how infrequently they are covered in sources. These information gaps can limit opportunities to improve the transport system. Our design – through its evidence requirements and metadata standards − will help fill these gaps, by prompting researchers and transport practitioners to generate much-needed new evidence from operational reports that aligns with research and evaluation reports.

We agreed 10 recommendations with our project steering group. The recommendations emphasise engaging CETI users early and sustaining their engagement. We consider that the breadth and quality of the CETI's content is critical to its success, but achieving and maintaining this will require ongoing investment and oversight. The CETI will provide a home for evidence, but it will also contribute to improving new evidence about transport interventions in New Zealand − it will set standards for what transport interventions are evaluated on and reported against, which will address a significant information gap.

References

Cleland, C. L., Jones, S., Moeinaddini, M., Weir, H., Kee, F., Barry, J., Longo, A., McKeown, G., Garcia, L., & Hunter, R. F. (2023). Complex interventions to reduce car use and change travel behaviour: An umbrella review. *Journal of Transport & Health*, *31*, 101652. https://doi.org/10.1016/j.jth.2023.101652

Hay, D. C. (2010). *Data Model Patterns: A Metadata Map*. Elsevier.

Ivory, V. (2023). *Reducing greenhouse gas emissions in communities–evidence and opportunities for change in Aotearoa* (Research Report ER80[2023]). BRANZ. https://www.branz.co.nz/documents/4363/ER80_-_Reducing_greenhouse_gas_emissions_in_communities.pdf

Ivory, V., Morton, S., Schmidt, J., Kingi, T. K., & Atatoa-Carr, P. (2013). Chapter 22 - Capturing the Capitals: A Heuristic for Measuring the "Wealth" of New Zealand Children in the Twenty-First Century: An Application to the Growing Up in New Zealand Longitudinal Cohort. In C. Banwell, S. Ulijaszek, & J. Dixon (Eds.), *When Culture Impacts Health* (pp. 251–268). Academic Press. https://doi.org/10.1016/B978-0-12-415921-1.00022-1

Mackie, H., Hawley, G., Scott, R., & Woodward, A. (2017). *Towards a safe system approach for cycling. Development and application of a cycling safety system model. Preparing New Zealanders for utility cycling* (Research Report 606). Waka Kotahi NZ Transport Agency. http://www.nzta.govt.nz/resources/research/reports/606/

Marewski, J. N., & Gigerenzer, G. (2012). Heuristic decision making in medicine. *Dialogues in Clinical Neuroscience*, *14*(1), 77–89. PMC.

May, A. D., Khreis, H., & Mullen, C. (2018). Option generation for policy measures and packages: An assessment of the KonSULT knowledgebase. *Case Studies on Transport Policy*, *6*(3), 311–318.

May, A. D., & Taylor, M. A. P. (2002). KonSULT-Developing an International Knowledgebase on Urban Transport Policy Instruments. *Australasian Transport Research Forum*. Australasian Transport Research Forum, Canberra, Australian Capital Territory.

Ministry for the Environment. (2022). *National Policy Statement on Urban Development—Updated 2022* (ME 1513). Ministry for the Environment. https://environment.govt.nz/publications/national-policy-statement-on-urban-development-2020-updated-may-2022/

Nan, Y., Ser, J. D., Walsh, S., Schönlieb, C., Roberts, M., Selby, I., Howard, K., Owen, J., Neville, J., Guiot, J., Ernst, B., Pastor, A., Alberich-Bayarri, A., Menzel, M. I., Walsh, S., Vos, W., Flerin, N., Charbonnier, J.-P., van Rikxoort, E., … Yang, G. (2022). Data harmonisation for information fusion in digital healthcare: A state-of-the-art systematic review, meta-analysis and future research directions. *Information Fusion*, *82*, 99–122. https://doi.org/10.1016/j.inffus.2022.01.001

Ogilvie, D., Bauman, A., Foley, L., Guell, C., Humphreys, D., & Panter, J. (2020). Making sense of the evidence in population health intervention research: Building a dry stone wall. *BMJ Global Health*, *5*(12), e004017. https://doi.org/10.1136/bmjgh-2020-004017

Pacheco, E., & Ivory, V. (2023). *A Narrative Literature Review of the Effectiveness of Interventions to Reduce Light Vehicle Travel* (Research Report 707; Waka Kotahi NZ Transport Agency Research Report). Waka Kotahi NZ Transport Agency. https://nzta.govt.nz/resources/research/reports/707

Panter, J., Guell, C., Humphreys, D., & Ogilvie, D. (2019). Can changing the physical environment promote walking and cycling? A systematic review of what works and how. *Health & Place*, *58*, 102161. https://doi.org/10.1016/j.healthplace.2019.102161

Rees, D., Masari, E., & Appleton-Dyer, S. (2020). *Transport impacts on wellbeing and liveability: Literature summary* (Research Report 669). Waka Kotahi NZ Transport Agency. https://www.nzta.govt.nz/assets/resources/research/reports/669/669-Transport-impacts-on-wellbeing-and-liveability.pdf

Shneiderman, B. (1998). Designing the user interface: Strategies for effective human-computer interaction. In *Designing the user interface: Strategies for effective human-computer interaction* (3rd ed.). Addison-Wesley. https://ui.adsabs.harvard.edu/abs/1998duis.book.....S

Sullivan, L. E. (2009). Heuristic. In L. E. Sullivan (Ed.), *The SAGE glossary of the social and behavioral sciences* (p. 234). Sage.

Tannahill, A. (2008). Beyond evidence—to ethics: A decision-making framework for health promotion, public health and health improvement†. *Health Promotion International*, *23*(4), 380–390. https://doi.org/10.1093/heapro/dan032

The Maryland Scientific Methods Scale (SMS). (2015, June 9). What Works Growth. https://whatworksgrowth.org/resource-library/the-maryland-scientific-methods-scale-sms/

Thomas, J. A., Malcolm, L. A., Cooper, D., Cross, T., Cai, L., Thomas, F., & Frith, B. (2022). *Safety interventions and their contribution to mode shift* (Research Report 701). Waka Kotahi NZ Transport Agency. https://trid.trb.org/view/2112871

Appendix A: Example sources and definitions used in the prototype catalogue of evidence

Appendix A is an Excel spreadsheet, available at www.nzta.govt.nz/resources/research/reports/725

The spreadsheet contains the catalogue's framework − its categories, filters and values, and their associated definitions and descriptions (see the 'Definitions' tab). The spreadsheet is populated with information we extracted from a small sample of sources, and our indicative assessments and judgements (see the 'Sources' tab). The content in the Sources tab is intended solely to develop the prototype and demonstrate its capability; it should not be used for decision-making purposes.

Transport Intervention Evidence