



Getting there from here: a literature review on vocational education and training reform implementation

Katherine Marie Caves, Severin Baumann & Ursula Renold

To cite this article: Katherine Marie Caves, Severin Baumann & Ursula Renold (2021) Getting there from here: a literature review on vocational education and training reform implementation, *Journal of Vocational Education & Training*, 73:1, 95-126, DOI: [10.1080/13636820.2019.1698643](https://doi.org/10.1080/13636820.2019.1698643)

To link to this article: <https://doi.org/10.1080/13636820.2019.1698643>



Published online: 04 Dec 2019.



Submit your article to this journal [↗](#)



Article views: 2218



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 8 View citing articles [↗](#)



Getting there from here: a literature review on vocational education and training reform implementation

Katherine Marie Caves , Severin Baumann and Ursula Renold 

KOF Swiss Economic Institute, ETH Zurich, Zurich, Switzerland

ABSTRACT

Vocational education and training (VET) reform is increasingly common, but there is more evidence on designing VET reforms than implementing them. This literature review searches the existing literature on VET reform implementation for key determinants, trends, and gaps using a determinant framework. We review 1,835 sources, coding 177 for 1,538 data points. We find that, like other implementation domains, resource- and stakeholder-related determinants are very important. For VET specifically stakeholders include employment-system actors and the relationships among actors are key. The major determinants are more like necessary than sufficient conditions, and mainly operation in conjunction with others. Europe is more represented than other continents. The findings in the literature are consistent over time, type, continent, and development status, but it is not clear if that is due to consensus or stagnation. The field is growing, however, so future research can develop theory by developing and testing hypotheses.

ARTICLE HISTORY

Received 1 February 2019

Accepted 11 November 2019

KEYWORDS

VET; implementation; VET reform; reform implementation

1. Introduction

Vocational education and training (VET) reforms are a major global topic, driven by new understandings of VET (e.g. Rageth and Renold 2019) and youth employment challenges (e.g. Pusterla 2016). There are decades of research on what constitutes good VET (e.g. Wolter and Ryan 2011), but much less on implementing it (Viennet and Pont 2017). This literature review synthesises the available evidence on reform implementation in VET to summarise the field, and potential future directions.

Implementation is the level to which a measure or policy is taken up in practice (Nilsen 2015). According to Li and Pilz (2017), 'implementation research focuses on the discrepancy between the way a measure has been planned and how it is implemented in practice, and is designed to gather information about the extent of uptake in practice and the factors involved' (p. 472).

We develop a determinant framework (Nilsen 2015) that summarises what might matter for implementing VET reforms. Using that framework, we synthesise

all available literature on VET reform implementation to answer the following questions: What are the commonly cited determinants of implementation success? How do different subgroups of literature differ? How have key determinants changed over time?

We expect that what matters for implementing VET is similar overall to what matters for implementing any policy reform, especially education reforms. We also expect that the differentiating characteristics of VET will also differentiate what matters for implementing VET reforms. Specifically, employers' involvement in VET (Bolli et al. 2018) will give them a role in implementing VET, unlike general education. In addition, VET programmes need to have different structures that balance actors' incentives (Wolter and Ryan 2011; Wolter, Mühlemann, and Schweri 2006) and need to update quickly to stay aligned with the labour market (Wallenborn 2010).

In addition to the general determinants of implementing VET reforms, we expect to find that the literature has gaps and blind spots. We also expect to find that there are trends over time and source types that reflect both growing understanding and general trends of popular topics. We discuss these patterns and the contributions that future VET implementation research can make.

2. Existing implementation frameworks

Nilsen (2015) differentiates five types of implementation models. *Process models* are basic, specifying only steps or phases. *Determinant frameworks* explain different implementation outcomes by specifying success factors and barriers. *Classic theories* originate from the substantive field of the policy itself (in this case VET), and can be applied alone or in conjunction with other models. For example, a classic theory can inform the determinants in the framework. *Implementation theories* are developed specifically to explain implementation. Finally, *evaluation frameworks* identify the criteria for successful implementation.

We develop a determinant framework influenced by the domains of policy implementation (e.g. Nilsen 2015; Winter 2012), general education reform implementation (i.e. Fullan 2015; Honig 2006; Kohoutek, 2013; Viennet and Pont 2017), and VET reform implementation (Fluitman 1999; Sultana 2008; Atchoarena and Grootings 2009). Determinant frameworks are not as sophisticated as implementation theories, are often overly generic, and do not always capture determinant interactions (Nilsen 2015). However, they are a step towards a domain-specific implementation theory and meaningful evaluation frameworks.

2.1 Implementation research

Implementation research originated in the early 1970s (Pressman and Wildavsky 1973). Its evolution is often described in three generations (see Honig 2006; Winter 2012). The first generation is mostly process models, or exploratory case

studies exploring gaps like design-implementation mismatch (Nilsen et al. 2013). These studies tend to view the policy process as linear, top-down, and evolving in discrete stages based on process models (*ibid*). This research is generally more concerned with describing the implementation process than predicting policy outcomes (Schofield 2001).

The second generation of implementation research focuses on determinant frameworks and implementation theories – building predictive models and understanding what factors influence the implementation process (Schofield 2001). Some models emphasise central planners in a hierarchical and administrative process, while others stress local context and the implementers themselves (Nilsen et al. 2013). This gave rise to the familiar top-down/bottom-up dichotomy, as well as efforts to integrate the two (Winter 1990; Goggin et al. 1990; Matland 1995). Over time, the integrated approaches became the accepted heuristic tool (Viennet and Pont 2017). One contribution of this research is the grouping of determinants into clusters (Najam 1995).

Goggin et al. (1990) strongly influence the third generation of implementation research, emphasising implementation theories and evaluation frameworks through rigorous and longitudinal research designs, quantitative methods, and theory-driven hypothesis testing. Research from this era also takes a comparative and multi-theoretical approach (Schofield 2001). However, this paradigm can have unrealistic methodological requirements (O'Toole 2000; Winter 2012), which makes it unfeasible for the present purpose. Partial theory testing is especially useful because the implementation-research field is still populated by a multitude of theoretical frameworks and models, none of which has become the gold standard (Nilsen 2015).

Our review is most aligned with the second generation of implementation research and its determinant frameworks. One of the most widely applied and generally applicable frameworks from that period is the 5C framework (Najam 1995), which describes implementation processes through five causal clusters of determinants. We choose this framework because it synthesises a great deal of implementation research without focusing on a specific context or domain, allowing us to apply classic theories from VET reforms. Because of this domain ambivalence, the 5C framework is still widely used in a number of implementation contexts. We use the 5C model as the starting point for developing our determinant framework for implementing VET reforms.

Najam (1995) defines the five categories as follows:

- “The *Content* of the policy itself – What it sets out to do (i.e. goals); how it problematises the issue (i.e. causal theory); how it aims to solve the perceived problem (i.e. methods).
- “The nature of the institutional *Context* – The corridor (often structured as standard operating procedures) through which policy must travel, and by whose boundaries it is limited, in the process of implementation.

- “The *Commitment* of those entrusted with carrying out the implementation at various levels to the goals, causal theory, and methods of the policy.
- “The administrative *Capacity* of implementers to carry out the changes desired of them.
- ‘The support of *Clients and Coalitions* whose interests are enhanced or threatened by the policy, and the strategies they employ in strengthening or deflecting its implementation.’

We draw on the literature of general-education reform and VET reform for the classic theories we can apply to the 5C framework.

2.2 Implementation in general education

The current state of implementation research in education is similar to that of implementation research in general. It considers policy-making and implementation to be non-linear, dynamic, and integrated processes rather than products of centrally planned design (Gornitzka, Kyvik, and Stensaker 2005). Scholars have shifted from trying to reduce complexity to confronting and embracing it (Honig 2006), giving rise to new approaches like Complexity Theory, Organisational Theory, and Network Theory (Gornitzka, Kyvik, and Stensaker 2005; Honig 2009). These all highlight the importance of who is involved in the reform and their relationships to the project. Another trend is single-case studies, usually focused on interesting policy approaches or situations with specific resource conditions (e.g. Souto-Otero 2011).

The relatively large implementation literature in general education is our first source of potential determinants. An early pivotal work in this field is Cerych and Sabatier (1986) book, which identifies five critical implementation factors: a sound theory of action, unambiguous policy directives with adequate resources and coordination, management and political skills, support from key stakeholders at all levels, and a context where the policy is not undermined by changing conditions. Honig (2006) argues that educational reform depends on three broad factor clusters: policy, people, and places. Fullan (2015) and his co-authors are a major presence in this literature, focusing on school-level change, capacity, teacher ownership, and school leadership. They enumerate nine elements for successful implementation in three clusters: change characteristics (need, clarity, complexity, and quality/practicality), local characteristics (district, community, principal, teacher), and external factors (government and other agencies). Barber, Moffit, and Kihn (2010) model emphasises iteration, learning, and evaluation. The OECD (2011) focuses on institutions, resources, and a formal legal framework for the reform. Viennet and Pont (2017) summarise the literature on education policy implementation into a four-dimensional framework of smart policy design, inclusive stakeholder engagement, conducive context and coherent implementation strategy.

2.3 Implementation in VET

In addition to the determinants that we draw from general education implementation, some are specific to VET reform. Fluitman (1999) was one of the first to focus on VET reforms. He concludes that the obstacles to VET reforms are universal and well known from policy implementation research: lack of resources (monetary, informational, and human), lack of expertise and commitment, and administrative or regulatory constraints. Sultana (2008) develops a framework specifically for VET implementation, starting with the policy/people/places-framework (Honig 2006) plus a pace dimension. For VET, 'people' includes employers, their associations, and trade unions. OECD (2009) develops a model and typology of system-wide innovation in VET, highlighting innovation drivers, enablers, and barriers. According to that study, innovative VET systems need to be coherent and targeted systems designed to sustain innovation, with set national priorities, formalised knowledge bases and research programmes, and stakeholders engaged in the VET policy dialogue.

The many comparative case studies of VET reform implementation are typically descriptive rather than theoretical. In contrast, Wilson (1993) examined eight countries, finding that developing-country VET reforms should be 'in concert with industrialisation, or industrial restructuring' to succeed (p.280). Interestingly, even this relatively informal conclusion is VET-specific, rather than derived from the general education implementation literature as suggested by Fluitman (1999).

2.4 Determinant framework for VET implementation

Table 1 summarises our determinant framework with the main citation for each determinant. Each determinant is phrased in terms of what the studies just reviewed consider helpful, for example as political will instead of opposition. The framework combines the education and VET implementation literatures, yielding a set of determinants we can search for in the empirical literature on VET reform implementation projects. Determinant frameworks do not capture interactions or specify theory (Nilsen 2015), making this framework a surveying tool for the literature, not a theory of action.

Under the *Content* dimension, we have six determinants. Content covers the characteristics of the reform being implemented, including the overall strategy or plan, accountability measures used to evaluate and ensure that actors achieve their intended outcomes, the presence of piloting or iteration in the reform process, the pace of the reform, its direction in terms of bottom-up or top-down, and its scope in terms of incremental or radical. For pacing, direction, and scope, we had to choose one end of each continuum to report whether a given study states that the determinant is helpful or hurtful – for example, pacing itself is neither helpful nor hurtful but a slow pace can be.



Table 1. Determinant framework by main source.

Cat.	Short name	Determinant description (vs. opposite)	Selected source frameworks
Content	Strategy	Clarity/strategy/vision (vs. confusion/short-termism/un-clarity)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017; Sultana 2008
	Accountability	Quality assurance/regulation/accountability (vs. lack thereof)	Viennet and Pont 2017; Fullan 2015
	Piloting	Piloting/iteration (vs. lack thereof)	Barber, Moffitt, and Kihn 2010
	Slow Pace	Slow pace (vs. fast pace)	Viennet and Pont 2017; Sultana 2008
	Bottom-Up	Bottom-up (vs. top-down)	Cerych and Sabatier 1986; Gornitzka, Kyvik, and Stensaker 2005
	Incremental	Incremental scope of change (vs. radical scope)	Barber, Moffitt, and Kihn 2010; Sultana 2008
	Coordination	Coordination/efficiency/management (vs. bureaucracy)	Cerych and Sabatier 1986; Sultana 2008; Fluitman 1999
	Context Fit	Context/institutional fit (vs. mismatch/bad fit)	Viennet and Pont 2017; Sultana 2008; Wilson 1993
	Decentralise	Decentralisation (vs. centralisation)	Viennet and Pont 2017
	Formalism	Formalism/legislation (vs. informality, no laws)	OECD 2011; Ryan 2000
Commitment	Strong Econ	Strong economy/low unemployment (vs. weaker)	Viennet and Pont 2017; Wilson 1993
	Ed Quality	Good existing education quality (vs. poorer)	Viennet and Pont 2017
	Political Will	Political will/demand-side (vs. disinterest/opposition)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017; Fullan 2015
	Cooperation	Cooperation/consensus (vs. conflict)	Viennet and Pont 2017; Sultana 2008; Wolter and Ryan 2011; Bolli et al. 2018; Rageth and Renold 2019
	Foreign Aid	Foreign assistance/influence (vs. lack thereof)	Viennet and Pont 2017
Capacity	Ownership	Ownership (vs. lack thereof)	Fullan 2015
	Low Turnover	Long political cycles/low turnover (vs. shorter/higher)	Fullan 2015
	Personnel	Personnel/expertise (vs. lack thereof)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017; Fullan 2015; Barber, Moffitt, and Kihn 2010; Sultana 2008; Fluitman 1999
Clients	Finances	Financial/material resources (vs. lack thereof)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017; Fullan 2015; Sultana 2008; Fluitman 1999
	Research	Data/research/evaluation (vs. lack thereof)	Cerych and Sabatier 1986; Viennet and Pont 2017; Fluitman 1999
	Time	Time (vs. lack thereof)	Viennet and Pont 2017; Fullan 2015; Sultana 2008
	Leadership	Strong/good leadership (vs. lack thereof)	Viennet and Pont 2017; Fullan 2015
	Employers	Employer engagement (vs. lack thereof)	Sultana 2008; Wilson 1993; Wolter and Ryan 2011; Bolli et al. 2018; Rageth and Renold 2019
	Educators	Educator engagement (vs. lack thereof)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017; Sultana 2008
	Intermediaries	Intermediary engagement (vs. lack thereof)	Honig 2006; Viennet and Pont 2017; Sultana 2008; Wolter and Ryan 2011; Bolli et al. 2018; Rageth and Renold 2019
	Trade Unions	Trade union engagement (vs. lack thereof)	Viennet and Pont 2017; Sultana 2008; Wolter and Ryan 2011; Bolli et al. 2018; Rageth and Renold 2019
	Community	Community/parent/student engagement (vs. lack thereof)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017; Sultana 2008; Fullan 2015
	Low Level	Low/local-level engagement (vs. lack thereof)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017
High Level	High/national-level engagement (vs. lack thereof)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017	
Mid Level	Mid/regional-level engagement (vs. lack thereof)	Cerych and Sabatier 1986; Honig 2006; Viennet and Pont 2017	

Under *Context*, there are also six determinants. Context reflects the time, place, situation, and existing education system. Coordination is when actors act efficiently together, with the opposite being bureaucracy. Context fit means the reform fits its environment. Under decentralisation, power structures are dispersed as opposed to the unified centralised model. In formalism, reform is sanctioned through laws and is part of the education system, as opposed to being informal or outside the legal and education systems. A 'strong' economy has industry and relatively low unemployment, so that industry and the labour market can support and absorb VET students and graduates. Finally, a higher-quality existing education system should prepare youth for VET and support their further development after secondary education.

Commitment has five determinants. Political will comes up frequently in the reviewed frameworks, and captures whether stakeholders welcome the reform, as opposed to disinterest or even opposition. Cooperation is the ability of stakeholders to work together, and its opposite is conflict. Foreign influence is especially relevant in developing contexts, and refers to outside advising, resources, or policy ideas in the project. Ownership is actors' and affected parties' sense that they have choice and control rather than imposition. Finally, low turnover refers to a relatively stable reform project team and government leadership.

Capacity has five determinants, each referring to a resource type. These include personnel, finances, research or information, time, and leadership.

Clients has the most dimensions at eight because it captures both the type and level of stakeholder involved. Types include employers, educators, intermediaries like industry or employers' associations, trade unions, and the community. Levels are described simply in terms of high, medium, and low-level actors so they can be flexibly applied across institutional contexts.

3. Method

This is a systematic literature review, meaning we use a standardised, structured and protocol-driven methodology (Jesson, Matheson, and Lacey 2011). The aim is to be as transparent and unbiased as possible by answering a clearly defined research question and following a clearly designed review process. We analyse the results using simple linear regressions, but do not perform meta-analysis because we lack the necessary data.

3.1 Search strategy

We do not restrict our search to peer-reviewed journals because much VET research is conducted by international organisations. To ensure quality, we restrict ourselves to reports and studies accessible in (a) electronic research databases on VET and general education, (b) academic databases on related topics, or (c) international organisations doing VET research (see Table A1

online¹). We systematically searched those databases and snowballed additional sources from key papers' references during the coding process.

The search syntax was based on the terms 'vocational education' (including both VET and TVET), plus any of 'implementation,' 'reform,' 'change,' or 'innovation.' In Boolean search notation, that is:

'Vocational education' OR VET OR TVET
AND
Implementation OR reform OR change OR innovation

Each database has a unique search interface, so we modified the search terms to fit, for example, whether it automatically searches full-texts or only titles. Finding the right search term for each index was an iterative process until we identified how to access the same underlying Boolean search terms across databases (see [Table A1](#)). Our goal is to capture as much of the literature as possible, so we had no date requirements and included the sources used to build our framework if they had an empirical element.

We conducted all final searches in January of 2018, yielding 1,835 sources. To decide whether to include a paper, we read its title, abstract, and conclusion. Selected sources are concerned with change in initial (upper-secondary) VET and explicitly deal with implementation. We are aware that researchers have different opinions on defining implementation, but for this study simply using the term qualified a paper for selection. Finally, only English-language papers were included. After screening the initial 1,835 results, we chose 175. Snowballing added another 44 sources and we dropped 40 during coding, for 177 fully coded sources.

3.2 Coding strategy

We used our determinant framework ([Table 1](#)) to code success factors and barriers to VET reform implementation. We refined the framework and definitions by having two coders independently code 20 randomly selected papers, then compare and resolve disagreements through discussion. We repeated this process four times until we had aligned our coding and felt we could accurately capture what our sources report as helping or hindering implementation.

The three possible codes were positive, negative, and mixed. A positive determinant helps implementation; a negative determinant hurts implementation. When determinants are conditional, neutral, or complicated, the code is mixed. For example, the coordination determinant is positive when a source states that coordinating helped implementation, when it states that bureaucracy hindered implementation, or when a lack of coordination hindered implementation. Coordination is negative if bureaucracy helped implementation, the lack of bureaucracy hurt, or the lack of cooperation helped. If a source states that coordination matters but not how, if it helped at one point and hurt at another, or was otherwise complex, it is mixed. If coordination is not mentioned,

we do not code it for that source. Each determinant can only be coded once per source.

Two authors independently coded all 177 sources. We resolved disagreements every 25 sources through intensive discussion, including going back to the source to re-read relevant passages. This was time-consuming, but minimised subjectivity and maximised rigour. Average interclass correlation coefficient (ICC), a measure used to assess interrater reliability, was 0.81 before resolution. After resolution, we agreed 100%. We also noted sources' publication year, literature type (peer-reviewed or not), and the country/countries described by the source.

4. Results

We coded 1,538 mentions of our 30 determinants. We expect almost entirely positive codes because the framework is designed to capture success factors. There are 1,414 positive codes, 43 negatives, and 81 mixed. [Table 2](#) summarises mentions for each determinant by dimension.

At the dimension level, Capacity has the highest average count (67 per determinant) followed by Clients (53.63) and Commitment (49.40), then Content (47.50) and Context (40.33). Capacity and Clients are also the most positive (98.21% and 95.80%, respectively), and Commitment, Context, and Content are all mainly positive (89.88%, 85.95%, and 85.61%, respectively). Context has the most mixed determinant mentions (5.37%), followed closely by Content (5.26%). Commitment has a few mixed mentions (3.24%) and Clients and Capacity have very few (1.17% and 0.60%, respectively). Finally, Content and Context have the most negative mentions (9.12% and 8.68%, respectively), with Commitment (6.88%) not far behind. Capacity and Clients have very few negative mentions of their determinants (1.19% and 3.03%, respectively).

Based on the code counts, we find three categories of determinant: key success factors that come up frequently and positively, success factors that come up less often but still positively, and open questions that have more mixed and negative codes. Key success factors are mentioned at least 50 times with at least 90% positive mentions. Success factors come up less but are still at least 90% positive. Determinants that need more research are the rest, with more mixed or even negative codes. [Figure 1](#) summarises our determinant framework results. The next section analyzes each determinant qualitatively, using illustrative quotations that highlight mainstream and interesting opposition views.

4.1 Results for content determinants

Strategy a key success factor. In general, a strong strategic plan, clarity, and a sense of vision are prerequisites for successful implementation (e.g. Grootings and Nielsen 2005).

Table 2. Determinant codes by dimension.

Category	Determinant	Positive	Mixed	Negative	Total
Content	Strategy	88	2	0	90
	Accountability	54	3	3	60
	Piloting	38	4	1	43
	Slow Pace	33	4	3	40
	Bottom-Up	18	10	2	30
	Incremental	13	3	6	22
Context	Coordination	73	1	0	74
	Context Fit	64	1	0	65
	Decentralise	14	11	9	34
	Formalism	39	7	3	49
	Strong Econ	13	1	1	15
	Ed Quality	5	0	0	5
Commitment	Political Will	73	3	0	76
	Cooperation	68	0	0	68
	Foreign Aid	35	13	8	56
	Ownership	32	1	0	33
	Low Turnover	14	0	0	14
Capacity	Personnel	95	0	0	95
	Finances	85	2	1	88
	Research	84	1	1	86
	Time	36	1	0	37
	Leadership	29	0	0	29
Clients	Employers	104	1	0	105
	Educators	68	4	1	73
	Intermediaries	81	1	2	84
	Trade Unions	38	0	0	38
	Community	17	0	0	17
	Low Level	43	4	0	47
	High Level	35	1	2	38
	Mid Level	25	2	0	27
Totals		1,414	81	43	1,538

Notes: Numbers are counts of determinant mentions in the sample of 177 sources.

Positive mentions indicate the determinant improves implementation probability, mixed indicates that it plays a conditional or complex role, and negative mentions imply the determinant reduces implementation probability.

Accountability is a key success factor. According to ETF (2017), ‘quality assurance links the other components of a qualification system – legislation, stakeholders and institutions’ (p109). Evaluation and accountability are key for implementation, especially as an enabling factor for other determinants like decentralisation (ETF 2014).

Bottom-Up needs further research. Some sources recommend a ‘complementary top-down and bottom-up approach’ (Viertel and Grootings 2001, 1). Although both approaches have worked, Sultana (2008) finds, ‘Some policies are better implemented in a top-down manner, while others are more likely to have staying power if they are incubated within the school environment itself’ (p19).

Piloting needs further research. Piloting should improve implementation by letting reforms experiment with new actors, models, and relationships (Stoica 2003). Oates (2008) blames lack of piloting for previous implementation failures, although piloting lets reformers start where they are. Pilot schools have both strengths and weaknesses according to Viertel and Grootings (2001), and Stoica (2003) notes they can be expensive.

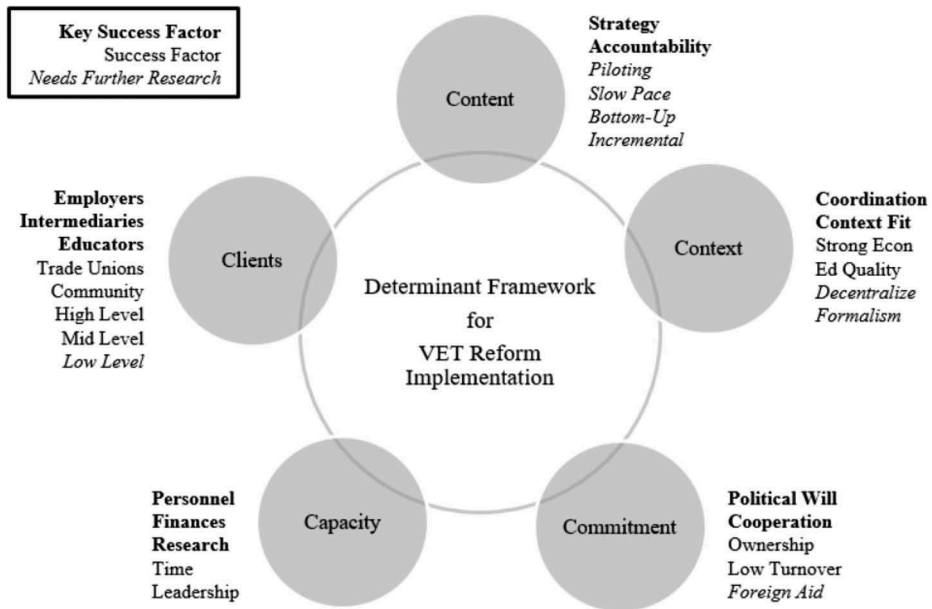


Figure 1. Determinant framework for VET reform implementation.

Slow Pace needs further research. Ertl (2000) and others caution VET projects against moving too slowly for employment actors, but allow that speed is fine when the plan is very clear. Mitchell et al. (2003) argue that fast and radical change are difficult for teachers, who may oppose reform. Finally, Oates (2008) and Cedefop (2009) argue that the problem is temporal discontinuity and that VET reforms are a multispeed process.

Incremental needs further research. Most sources found radical reforms overambitious (e.g. Hummelsheim and Baur 2014). Comyn and Barnaart (2010) blame failure on being too radical for the institutional structure.

4.2 Results for context determinants

Coordination is a key success factor. Cedefop (2017) states that the main implementation challenge is ‘getting organised’ (p25). Many sources find uncoordinated bureaucracy a barrier to reform (e.g. Castel-Branco 2008).

Context Fit is a key success factor. Hummelsheim and Baur (2014) argue that a VET model can only be successful if it ‘reflect[s] the existing conditions in the country’ and is ‘adapted to its unique social, cultural and economic objectives’ (p. 279). According to Hoppe, Burmester, and Ebben (2011), the ‘success of a strategy stands or falls with how well the strategy fits with other education sector policies and strategies’ (p.283). Essentially, ‘new policies need to be strategically linked to goals and outcomes for national education systems and must be firmly related to concrete national policy priorities as well as anchored up in specific country institutional contexts’ (ETF 2012, 9).

Strong Economy is a success factor. Kingombe (2011) states that ‘even the world’s most sophisticated and expensive programme is doomed to fail if the labour market cannot absorb the students, despite their skills and expectations’ (p61). In contrast, Cedefop (2012) finds that a strong labour market might hinder implementation by lowering political will and a weak labour market might encourage VET reform as a crisis-response measure, facilitating implementation as actors pull together to solve the problem. However, the same source finds weak labour markets have slowed down VET implementation due to economic uncertainty.

Education Quality is a success factor. Wallenborn (2010) argues that weak education quality and penetration prevented progress because there were no qualified trainers available. Other sources point out that stigma from a bad existing system attaches to reforms and prevents them from building momentum (Paik 2014).

Decentralise needs further research. Multiple sources advocated balancing the two extremes, with ETF (1999) recommending regionalism as the balancing method. Others argue for central coordination and oversight with instructional autonomy (Atchoarena and Delluc 2002) enabled by accountability (ETF 2014). ETF (2017) contradicts this, stating, ‘putting too many tasks under one roof can blur responsibilities between the agency and other actors ... Therefore, concentrating all related tasks in a single technical agency is not a feasible solution’ (p. 85).

Formalism needs further research. According to ETF (2017), ‘laws can be enablers, but can also create rigidities that only inhibit reform’ (p. 18). However, formal systems can be harder to implement (Castel-Branco 2008).

4.3 Results for commitment determinants

Political Will is a key success factor. Fluitman (1999) states, ‘perhaps the most serious constraint is people who resist change’ (p64). Will may not be enough for success without a good plan (Akyeampong 2005). Grossmann and Naanda (2006) find ‘it is political will rather than institutional capacity that is a key ingredient for successful reform’ (p16).

Cooperation is a key success factor. Well-designed cooperation mechanisms ‘where all stakeholders can exchange information, discuss problems and challenges, and develop ... solutions’ are key (Paik 2014, 31). ETF (2017) says no actor can implement alone. Arribas (2016) challenges the concept of cooperation as a pure input, framing engagement and interaction among actors as an outcome of implementation instead of its process.

Ownership is a success factor. It is usually positive, with one mixed mention when ownership was good but not important (Comyn and Barnaart 2010).

Low Turnover is a success factor. Reforms take time and planning, so changing governments can bring them to a halt and systems in flux are not able to implement reforms (ETF 2014).

Foreign Aid needs further research. The most successful international actor at implementing VET reform is the European Union, which uses soft policy, monitoring, and evaluation to drive reform (e.g. Arribas, 2016). When international help hurts implementation, it may be due to donors' uncoordination and conflict or, as Bartlett (2013) theorises, because of long chains of principals and agents.

4.4 Results for capacity determinants

Personnel is a key success factor. Grootings and Nielsen (2005) state, 'Teacher competences and teacher education are the key factors in all processes of VET reform anywhere' (p25). However, Mitchell et al. (2003) state, 'Innovation can occur without VET teachers being trained as innovators' (p101). ETF (2017) emphasises the need for 'professionals dealing with the implementation' (p85).

Finances is a key success factor. Overall, funding helps, including external resources (Wallenborn 2010). However, money must be appropriately directed: Atchoarena and Delluc (2002) find 'subsidies or incentives offered to employers ... have not produced the desired effect' (p13).

Research is a key success factor. Research helps implementation overall (e.g. Castel-Branco 2008). Different evidence types are useful in different reform contexts, and Sohn, Kang, and Lim (2017) state evidence is only useful if implementers know how to apply it. Ertl (2000) cautions against burying implementers in too much evidence.

Time is a success factor. Sources were frequently unclear on whether insufficient time or inadequate speed hindered a reform. Dorleans et al. (2011) illustrate this challenge: 'at this point, the question may be asked whether the pace of the reforms is slower than expected or whether, in fact, it is the time-frame that is unrealistic' (p. 14). However, one case ran out of time despite lasting two decades (Rekkor, Ümarik, and Loogma 2013). One source argues that neither time alone nor pace hindered implementation, but rather 'temporal discontinuity,' or lack of time plus improper sequencing of steps (Oates 2008).

Leadership is a success factor. Comyn and Barnaart (2010) find success 'in those institutions where managers and leaders fully embraced the reforms' (p.62). Bartlett (2013) says that an obstacle was lacking 'a "champion" of the reforms within the administration' (p. 342).

4.5 Results for clients determinants

Employers are a key success factor. ETF (2017) states, 'Stakeholders from the world of work must have a role, as a prerequisite for systemic change' (p19). Atchoarena and Delluc (2002) agree, saying about VET systems that 'experience has shown that such a system can hardly be effective if most employers oppose it' (p13). This determinant might be the most obvious differentiator of VET reform implementation from general education.

Intermediaries are a key success factor. According to OECD (2009), intermediaries contribute by acting to 'share, spread and diffuse innovations' and 'link the micro level ... with the macro level' (p.49). Intermediaries can help employers that lack capacity by reducing their administrative load and increasing efficiency (*ibid*).

Educators are a key success factor. Educators make sense of a change process, although teachers might resist a project if it lacks practical utility or impinges on their autonomy (Rekkor, Ümarik, and Loogma 2013). Educators need capacity and incentives, because 'innovation ... demands more effort and time than routine teaching, and without incentives there is a low chance of success' (ETF 2014, 53).

Trade Unions are a success factor. Kis and Field (2009) state that trade unions are important because they can 'constructively counterbalance the interest of employers' (p38). However, they can also play a negative role if they have 'incentives to reduce access to shortage occupations, to maintain wages and union bargaining power' (p38).

High Level actors are a success factor. High-level actors are useful for legislation and regulation (e.g. Grootings and Nielsen 2005).

Mid Level actors are a success factor. ETF (1999) argues that the meso-level is the solution to the question of centralisation or decentralisation, arguing that each of the extremes comes with its own problems.

Community is a success factor. Grossmann and Naanda (2006) write that low involvement of trainees in the reform process is dangerous for implementation, often increasing trainees' resistance to change (p.42–43). Parents are powerful stakeholders and involving them in the reform fosters a sense of ownership in the community, improving sustainability (ETF 2012).

Low Level needs further research. A study by Powell (2001) comparing the implementation of TVET projects in Jamaica and Gambia finds local involvement especially important when a project is initiated by a foreign aid agency, as it guarantees that local context is taken into account and creates a sense of ownership in the recipient country.

5. Discussion

Our results show how often framework determinants appear in the literature, but not necessarily whether there is bias in the literature itself. We address this through subsample analysis, looking for trends over time, between source types, in different locations, and in different development contexts. In addition, it is not clear how our results relate to a theory of VET implementation. We address this in two ways, first discussing potential interactions among determinants using covariance analysis quantitatively and sources' comments qualitatively, and second by critically considering when scholars might discuss certain topics more than others.

5.1 *Subsample analyses*

The most straightforward comparison for subsamples of the literature is simple linear regression of total mentions by determinant. [Table 3](#) shows the comparisons by source age, type, and the development status of the subject country or countries. R^2 values show the degree of similarity between groups, with values closer to zero more different and those close to one very similar. To find groups' distinctive determinants, we show the standard residual for each. Standard residuals are the difference between actual and expected values, divided by the standard deviation of expected values. We highlight those that are or significantly different at $p < 0.05$ and $p < 0.1$.

5.1.1. *Source age*

Publication years range from 1984 to 2017, skewing towards more recent literature with a mean year of 2006 and a median of 2009. We divided the sample into approximate halves before 2009 and from 2009 onwards, with 87 sources in the former and 90 in the latter group. The groups are very similar ($R^2 = 0.87$), shown in [Table 3](#). There are no significantly different determinants between the two groups of sources.

The similarity between older and newer sources can mean either that the field has achieved consensus or that it has not progressed. Because there are hardly any reviews or frameworks related to VET reform implementation, there have been few opportunities to collect into theory. The skew towards newer literature indicates that the field is growing.

5.1.2. *Source type*

Of the 177 coded sources, 62 are peer-reviewed literature, and 115 are non-peer-reviewed sources including 49 books and 66 others – mostly policy reports from international agencies. There were more non-peer-reviewed sources so there are more mentions overall in non-peer-reviewed (1,116) than peer-reviewed (422) literature. Therefore, policy reports and books have more power in our results than journal articles. Some authors, notably Pilz, Fluitman, Oates, and Grootings, appear in both source types. However, the two types are not very similar ($R^2 = 0.49$), so the greater presence of non-peer-reviewed sources is not trivial.

Non-peer-reviewed literature mentions intermediaries and employers more often than peer-reviewed sources, and time less often. These three determinants do not entirely explain the difference between groups, but they are the most important. With non-peer-reviewed literature being such an important part of the body of knowledge in this field, basic research needs to both reflect and inform that work.

Three organisations make up large parts of the non-peer-reviewed literature: ETF, Cedefop, and OECD. To check whether any institutional idiosyncrasies are driving results, we performed separate subsample analyses of each compared to



Table 3. Subsamples by source age, type, and development status.

Category	Determinant	Source Age				Source Type				Development Status			
		Pre-2009	2009 On	Standard Residual	Scholarly	Non-Scholarly	Standard Residual	Developed	Develop-ing	Standard Residual			
		47	43	-0.40	27	63	0.34	29	52	1.35			
Content	Strategy	29	31	0.02	11	49	1.16	27	26	-0.75			
	Accountability	25	18	-0.66	14	29	-0.56	16	25	0.10			
	Piloting	22	18	-0.44	16	24	-1.14	17	19	-0.51			
	Slow Pace	15	15	-0.15	10	20	-0.74	11	15	-0.34			
	Bottom-Up	11	11	-0.16	6	16	-0.57	8	14	-0.17			
	Incremental	38	36	-0.26	25	49	-0.41	21	46	1.51			
	Coordination	33	32	-0.20	20	45	-0.12	23	30	-0.06			
	Context Fit	14	20	0.30	7	27	0.08	8	20	0.35			
	Decentralise	23	26	0.09	14	35	-0.15	13	31	0.88			
	Formalism	11	4	-0.69	5	10	-0.87	2	12	0.17			
Context	Strong Econ	1	4	0.05	2	3	-1.02	1	3	-0.53			
	Ed Quality	38	38	-0.11	28	48	-0.81	29	38	0.12			
	Political Will	31	37	0.33	16	52	0.81	26	34	0.03			
	Cooperation	29	27	-0.28	13	43	0.52	14	36	1.23			
	Foreign Aid	13	20	0.37	13	20	-1.08	14	14	-0.69			
	Ownership	7	7	-0.17	2	12	-0.40	2	10	-0.01			
	Low Turnover	47	48	-0.02	34	61	-0.58	33	53	1.09			
	Personnel	50	38	-0.99	31	57	-0.52	30	51	1.17			
	Finances	39	47	0.49	23	63	0.79	40	39	-0.73			
	Research	19	18	-0.22	18	19	-1.71*	21	15	-1.20			
Capacity	Time	12	17	0.22	10	19	-0.81	12	15	-0.43			
	Leadership	47	58	0.74	22	83	2.29**	46	47	-0.54			
	Employers	31	42	0.71	14	59	1.52	32	31	-0.74			
	Educators	34	50	1.09	11	73	2.83**	36	38	-0.47			
	Intermediaries	14	24	0.60	6	32	0.54	15	17	-0.51			
	Trade Unions	8	9	-0.09	1	16	-0.01	5	10	-0.27			
	Community	24	23	-0.21	10	37	0.44	15	28	0.45			
	Low Level	17	21	0.15	10	28	-0.18	15	19	-0.34			
	High Level	13	14	-0.08	3	24	0.32	10	16	-0.17			
	Mid Level	679	735	R² = 0.87	422	1116	R² = 0.49	571	804	R² = 0.66			

Notes: Standardised residuals are the difference between expected and actual values divided by the standard deviation of expected values. Higher standard residuals indicate that the determinant is more common than expected in the second/right-hand/specified group.

* Significant at $p < 0.1$, or > 1.64 standard deviations from the mean ** Significant at $p < 0.05$, or > 2 standard deviations from the mean.

the rest of the non-peer-reviewed literature (see [Table A2](#) online). ETF is most similar to all other non-peer-reviewed sources ($R^2 = 0.84$), with not significantly different determinants. Cedefop is broadly similar to the rest of the non-peer-reviewed literature ($R^2 = 0.63$), and is more likely to mention context fit, political will, and community engagement. OECD has the most sources and is the least similar to the other non-peer-reviewed literature ($R^2 = 0.45$), mentioning context fit and political will much less often but employers and intermediaries much more ($p < 0.05$). It also mentions foreign aid, personnel, and low-level actors less often ($p < 0.1$) than similar literature. Many OECD studies are part of *Learning for Jobs* (OECD 2011), which may drive these results.

In the non-peer-reviewed literature, sources often have a specific purpose or audience in mind while they report the progress of an implementation project. This affects the report's content, for example, a report for a ministry of education will emphasise contact with employers while one for the ministry of labour will emphasise connection with the education system. We are missing non-English-speaking VET research organisations like Korea's KRIVET and Switzerland's SFIVET, which publish mainly in their countries' native languages.

5.1.3. Development status

We find 81 sources each dealing with developed and developing countries, plus 15 sources that include multiple countries of mixed development status. Sources in developed and developing contexts are generally similar ($R^2 = 0.66$), with no significant differences. Despite having the same number of sources total, those on developed countries attribute outcomes to fewer determinants than those in developing countries (571 and 804, respectively) and developed sources mention fewer determinants as mixed factors. Developing countries are featured more often in non-peer-reviewed literature.

Developing sources mention Context factors more often, and developed sources are more likely to cite Client factors. This implies that there may be an order or priority list to which determinants matter, with specific actors mattering only once environmental factors are under control. It could also be that sources on developing countries are more likely to be written by outsiders, so they are more concerned with understanding the context.

Because European countries are dominant in the sample, the difference between developed and developing sources may actually come from a continent difference. However, we have a number of developing European countries in the sample thanks to post-Soviet reform efforts in Eastern European countries. Comparing developing and developed countries within Europe, we see a similar pattern to the overall development-status comparison. European developing countries are broadly similar to its developed countries ($R^2 = 0.68$), with only time different between groups as it comes up more in developed countries. The difference between developed and developing countries does not appear to arise from a difference between Europe and the world.

5.1.4. *Continents*

The biggest bias in the literature is towards European countries, which make up 50% of our sources. Of the rest, 16% are multi-continent, 11% deal with Asian countries, 10% African, 6% Oceanic, 5% North American, and 2% South American. However, European sources are very similar to all other sources ($R^2 = 0.82$), with no significantly different determinants. Across continents, the general pattern of mentions is strikingly similar (see [Table A3](#) online).

Multi-continent sources strongly agree with the general pattern ($R^2 = 0.80$). These sources mention context fit more often than all other sources ($p < 0.1$), possibly because they compare across contexts. Asian-country sources are similar to others ($R^2 = 0.72$), mentioning coordination more than others ($p < 0.1$). Most of our Asian sources are in Korea and China, and tend to be top-down reforms. African-country sources are somewhat similar to the others ($R^2 = 0.62$), with more mentions of personnel and finances ($p < 0.1$). This could be due to greater poverty, or due to less experience in African countries with formal VET systems, requiring more start-up capital.

We do not discuss Oceania or North and South America because there are not enough sources. There are three major explanations for the gap. First, our search is in English and there might be existing sources in other languages or using alternative terminologies. Second, VET may be less prevalent in the under-represented continents. Finally, VET could be a low priority for research or reform. Perhaps the most important contribution of this subsample is what it cannot tell us: implementation determinants in non-European countries, especially those in North and South America.

5.2 *Covariance analysis*

Determinant frameworks do not capture interactions. Covariance analysis shows which determinants appear often with others, partially addressing this issue. We use a simple covariance table (see [Table A4](#) online) of total mentions. Highly covarying determinants are usually the most frequent: intermediaries (0.042), financial resources (0.042), and coordination (0.041) are all key success factors. However, low-level actors (0.040) are highly covariant without overall high frequency, covarying (>0.05) with mid-level and high-level actors, accountability, context fit, decentralisation, cooperation, ownership, personnel, finances, research, and educators. The strongest covariances between two dimensions are employers-intermediaries (0.165), personnel-finances (0.117), educators-intermediaries (0.115), and employers-educators (0.111). The second is entirely capacity, but the other three combine actors from the education and employment systems.

None of the determinants in the key success factors group have low covariance (<0.02). The least covarying determinants are education quality (0.005), incremental scope (0.01), a strong economy (0.011), time (0.015), low turnover (0.017), and slow pacing (0.018).

Time is the only determinant with a negative average covariance, though it is weak. It is least likely to appear with intermediaries (-0.05) and employers (-0.03). This is difficult to interpret, so the relationship between time and implementation, mediated or moderated by other determinants, is worth further investigation.

5.3 Interpreting frequency

We cannot say that a determinant matters in practice, only that it comes up frequently in literature. Frequent determinants may be important, very broad, or simply fashionable. Some of the most commonly mentioned determinants seem straightforward – they appear in every implementation framework – but interpretation is difficult. For example, the crucial role of intermediary and employer stakeholders fits with VET classic theory, but may be overstated in a literature generally directed at education actors. Research is a key success factor – according to researchers. Popular concepts like political will can be very broad, and distinguishing between cooperation and coordination can become difficult in practice. This latter problem highlights the urgency of developing theory-based, operational definitions for all determinants. Finally, even the most often-cited success factors are probably necessary, not sufficient, conditions for implementation.

The consistency across the field is a good indication that key success factors' frequency is meaningful, but we should be wary of stagnation or repetition. Many of the researchers whose work we review here appear in more than one source, so they have experience that supports – but may also bias – their later work. If biased, certain determinants could show up simply because a researcher is accustomed to looking for them. However, the subsample analysis shows that this is unlikely to be driving the results in most cases. In this case, consistency is most likely a sign of consensus, not problems.

The determinants that need further research may be too narrowly defined, genuinely unimportant, or interdependent. Some controversies reflect changing trends, like foreign involvement in reforms becoming more nuanced after waves of policy-making, -advising, -borrowing, and outright transfer. It seems that international soft policy and capacity-building are helpful, while supply-side policy, contingent aid, and VET-as-export may create problems. Other unclear situations may be due to framework weaknesses or dependencies among determinants.

6. Conclusions

We hypothesised that VET implementation would be similar to overall and general-education implementation, with differentiation in employers' role and stakeholder interactions. We find that our framework broadly aligns with implementation in general, and that the VET-specific determinants like employers and intermediaries are key. Cooperation and coordination are both key success

factors, possibly more important in VET than general education because of the many stakeholder types.

We also expected to find gaps and blind spots in the literature, and a lack of systematic progression. We do find that non-European countries are underrepresented, though developed and developing countries are balanced. We find no significant changes in the literature over time, suggesting that the lack of structure may be causing some stagnation. Generally, we are missing a strong theoretical approach to explaining implementation success and failure. However, the field is growing.

The most important determinants of VET reform implementation are resources and stakeholders. Personnel, financial, and research resources are key along with employer, intermediary, and educator stakeholders. Characteristics of the reform and its context are also important, especially the strategy itself, coordination, political will, cooperation, and context fit. Some determinants, especially in the areas of reform context and actors' interactions, are unclear. Decentralisation, foreign assistance, reform scope, pacing, and direction (bottom-up or top-down) are among these conflicted factors.

When factors are important, they are more like necessary than sufficient conditions, and tend to work together rather than alone. The most important determinants are also the most interconnected, especially among stakeholders and resources. Stakeholders frequently appear in combinations that link education and employment. Coordination seems to combine with other success factors like financial resources and a legal framework. Low-level actors initially appear unimportant, but are one of the most strongly interconnected determinants.

6.1 Limitations

We have four main sources of limitations: the limitations of the literature itself, our determinant framework, content analysis and our coding process, and our analysis. First, the literature and our sample of sources can create biases. Our sample is dominated by European-continent sources and by non-peer-reviewed literature, even if the subsample analysis shows that these groups do not drive the results. We search only in English, and use VET and TVET search terms instead of specific country-level programmes. We are as systematic as possible in our search and sampling, but database ideosyncracies forced us to use multiple search approaches.

The determinant framework prioritises feasibility and completeness over context and interaction, and is likely to change in the future. Determinant frameworks do not account for items' interactions or relative importance. Identifying success factors and barriers faces does not specify a model of implementation (Gornitzka, Kyvik, and Stensaker 2005) or account for context (OECD 2009). In practice, there should be 'synergistic effects' (Nilsen 2015, 5) among determinants (Winter 2012). In addition, our goal was to make the

determinant framework mutually exclusive and collectively exhaustive, but there are already signs that we will drop, reorganise, add, or split determinants upon further testing.

Coding and content analysis enables us to process a huge amount of data, but our definitions are relatively weak and coding can be reductive. We often relied on authors' apparent usage definitions (Popper 2014) for both determinants and implementation success. Therefore, we are inaccurate if sources have inconsistent definitions. In addition, our codes are very simple – valence, not strength or impact – and we capture at most one mention per determinant per paper. These do not capture interrelationships, such as Oates's (2008) claim that temporal discontinuity matters more than time or pace alone, or Akyeampong's (2005) that political will is insufficient without strategy. Because we only code one mention per paper and do not capture importance, we cannot differentiate strong from weak determinants within sources.

Finally, our analysis is descriptive, not causal. We are indifferent to reforms' successful implementation or their impact. Successful implementation means the planned reform happened, while a successful reform, once implemented, should have impact (Viennet and Pont 2017). Although reforms' objectives and policy tools affect the implementation process (Honig 2009), we do not differentiate reform types. We do not analyse, for example, how reforms that build a VET system from scratch are different from those improving an existing system.

6.2 Final thoughts and future directions

The challenge for future research is building a theory of VET reform implementation. Systematically consolidating our empirical knowledge base is a start, and theory can develop testable hypotheses. Context dependence and interdependence is a key issue, as is the question of what it means for a VET reform to be successful. Finally, research in this field must be actionable and provide usable recommendations that increase the probability of successful implementation.

A number of interesting questions have emerged for further investigation. How do we define and measure each determinant? What makes each key success factor important? Why do low-level actors appear in conjunction with other factors so frequently, despite their low individual importance? Is stakeholder engagement most productive when it contributes to education-employment linkage? When is foreign aid or intervention helpful? What parts of the field's consistency are consensus, and which are stagnation? What can we learn from new geographies? How are implementation determinants affected by the reform's goals?

Despite its opacity, this field is one of surprising agreement. Future researchers can mine detailed case studies for new insights and capitalise on growth to develop new studies. There is a sense of urgency around VET, and enormous potential for well-designed and successfully implemented VET systems to impact young people, economies, and societies.

Note

1. <http://www.cemets.ethz.ch/research/research0/2019/online-appendices—getting-there-from-here-literature-review.html>.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Katherine Marie Caves  <http://orcid.org/0000-0002-3450-5942>

Ursula Renold  <http://orcid.org/0000-0003-4196-0019>

References

- Akyeampong, A. K. 2005. "Vocationalisation of Secondary Education in Ghana." In *Vocationalisation of Secondary Education Revisited*, edited by J. Lauglo and R. Maclean, 149–216. Dordrecht, Netherlands: Springer.
- Arribas, J. 2016. "Governance Dynamics and the Application of the Multilevel Governance Approach in Vocational Education and Training (VET) in the European Neighbourhood Countries: The Case of the ENPI South Region." *European Journal of Education* 51 (4): 495–512. doi:[10.1111/ejed.2016.51.issue-4](https://doi.org/10.1111/ejed.2016.51.issue-4).
- Atchoarena, D., and P. Grootings. 2009. "Overview: Changing National VET Systems through Reforms." In *International Handbook of Education for the Changing World of Work*, edited by R. Maclean and D. Wilson, 365–378. Dordrecht, Netherlands: Springer.
- Atchoarena, D., and A. M. Delluc. 2002. *Revisiting Technical and Vocational Education in sub-Saharan Africa: An Update on Trends, Innovations and Challenges*. Paris, France: UNESCO, IIEP.
- Barber, M., A. Moffit, and P. Kihn. 2010. *Deliverology 101: A Field Guide for Educational Leaders*. Thousand Oaks, CA: Corwin Press.
- Bartlett, W. 2013. "International Assistance Programmes and the Reform of Vocational Education in the Western Balkans: Sources of Policy Failure." *Southeastern Europe* 37 (3): 330–348. doi:[10.1163/18763332-03703005](https://doi.org/10.1163/18763332-03703005).
- Bolli, T., K. Caves, U. Renold, and J. Buergi. 2018. "Beyond Employer Engagement: Measuring Education-employment Linkage in Vocational education and Training Programmes." *Journal of Vocational Education and Training* 70 (4): 524–563.
- Castel-Branco, E. 2008. *Vocational Education and Training Challenges and Opportunities in the Southern Caucasus, Cross-Country Report, Armenia-Azerbaijan-Georgia*. Turin, Italy: European Training Foundation.
- Cedefop (2009). *Continuity, Consolidation and Change: Towards a European Era of Vocational Education and Training. Cedefop Reference Series 73*, Luxembourg: Publications Office of the European Union.
- Cedefop. (2012). "Curriculum Reform in Europe. The impact of learning outcomes." *Research papers* No. 29. Luxembourg: Publications Office of the European Union. doi:[10.1094/PDIS-11-11-0999-PDN](https://doi.org/10.1094/PDIS-11-11-0999-PDN).
- Cedefop. 2017. *Global Inventory of Regional and National Qualifications Frameworks 2017. Volume I: Thematic Chapters*. Luxembourg: Publications Office of the European Union.

- Cerych, L., and P. Sabatier. 1986. *Great Expectations and Mixed Performance: The Implementation of Higher Education Reforms in Education*. Stoke on Trent, England: Trentham Books.
- Comyn, P., and A. Barnaart. 2010. "TVET Reform in Chongqing: Big Steps on a Long March." *Research in Post-Compulsory Education* 15 (1): 49–65. doi:10.1080/13596740903565335.
- Dorleans, M., K. Holmes, J. Masson, S. Nielsen, D. Ouzoun, and H. Zelloth. 2011. *The Torino Process: Evidence-based Policy Making for Vocational Education and Training*. Turin, Italy: European Training Foundation.
- Ertl, H. 2000. "The Transition of Vocational Education and Training in Eastern Germany: Notes on the Role of European Union Programs." *Comparative Education Review* 44 (4): 464–492. doi:10.1086/447630.
- ETF. 1999. *Review of Progress in Vocational Education and Training Reform of the Candidate Countries for Accession to the European Union in the Light of Developments in European Policy on vocational training*. Turin, Italy: European Training Foundation.
- ETF. 2012. *ETF Yearbook 2012: Evaluation And Monitoring Of Vocational Education And Training Systems And The Role Of Evidence-Based Policy In Their Reforms*. Luxembourg: Publications Office of the European Union.
- ETF. (2014). *Torino Process 2014: A cross-country report - Moving skills forward together. Trends, challenges and successes in vocational education and training in the ETF partner countries*. Turin, Italy: European Training Foundation.
- ETF. 2017. *Getting Organised For Better Qualifications: A Toolkit*. Turin, Italy: European Training Foundation.
- Fluitman, F. 1999. "The Roots and Nature of Reforms in Vocational Education and Training: An Analytical Framework and Some Examples." *Prospects* 29 (1): 55–65. doi:10.1007/BF02736825.
- Fullan, M. 2015. *The New Meaning of Educational Change*. 5th ed. New York, NY: Teachers College Press.
- Goggin, M., A. Bowman, J. Lester, and L. O'Toole. 1990. *Implementation Theory and Practice: Toward a Third Generation*. Glenview, IL: Scott-Foresman.
- Gornitzka, Å., S. Kyvik, and B. Stensaker. 2005. "Implementation Analysis in Higher Education." In *Reform and Change in Higher Education: Analysing Policy Implementation*, edited by Å. Gornitzka., M. Kogan, and A. Amaral, 35–56. Dordrecht, Netherlands: Springer.
- Grootings, P., and S. Nielsen. 2005. *ETF Yearbook 2005: Teachers and Trainers. Professionals and Stakeholders in the Reform of Vocational Education and Training*. Luxembourg: Office for Official Publications of the European Communities.
- Grossmann, M., and R. Naanda. 2006. *Back to the Future? the Challenges of Reforming Vocational Education and Training (VET) Systems: A Critical Analysing of Namibia's Current VET Reform Coventry*. England: Skills, Knowledge and Organisational Performance (SKOPE) Publications. (Research Paper 64).
- Honig, M. I., Ed. 2006. *New Directions in Education Policy Implementation: Confronting Complexity*. Albany, NY: State University of New York.
- Honig, M. I. 2009. "What Works in Defining "What Works" in Educational Improvement: Lessons from Education Policy Implementation Research, Directions for Future research." In *Handbook of Education Policy Research*, edited by G. Sykes, B. Schneider, and D. N. Plank, 333–347. New York, NY: Routledge.
- Hoppe, M., J. Burmester, and J. Ebben. 2011. "The Development of a Strategy for Vocational Education and Training: Experiences from Montenegro." *Research in Comparative and International Education* 6 (3): 273–284. doi:10.2304/rcie.2011.6.3.273.

- Hummelsheim, S., and M. Baur. 2014. "The German Dual System of Initial Vocational Education and Training and Its Potential for Transfer to Asia." *Prospects* 44 (2): 279–296. doi:10.1007/s11125-014-9311-4.
- Jesson, J., L. Matheson, and F. M. Lacey. 2011. *Doing Your Literature Review: Traditional and Systematic Techniques*. London, England: Sage.
- Kingombe, C. (2011). "Lessons for Developing Countries from Experience with Technical and Vocational Education and Training." *Working Papers*, London, England: International Growth Centre (IGC).
- Kis, V., and S. Field. 2009. *Learning for Jobs - OECD Reviews of Vocational Education and Training. Chile: A First Report*. Paris, France: Organization for Economic Cooperation and Development (OECD).
- Kohoutek, J. 2013. "Three Decades of Implementation Research in Higher Education: Limitations and Prospects of Theory Development." *Higher Education Quarterly* 67 (1): 56–79.
- Li, J., and M. Pilz. 2017. "Modularisation in the German VET System: A Study of Policy Implementation." *Journal of Education and Work* 30 (5): 471–485. doi:10.1080/13639080.2016.1243233.
- Matland, R. E. 1995. "Synthesizing the Implementation Literature: The Ambiguity-conflict Model of Policy implementation." *Journal of Public Administration Research and Theory* 5 (2): 145–174.
- Mitchell, J., B. Clayton, J. Hedberg, and N. Paine. 2003. *Emerging Futures: Innovation in Teaching and Learning in VET*. Melbourne, Australia: Australian National Training Authority.
- Najam, A. (1995). "Learning from the Literature on Policy Implementation: A Synthesis Perspective." *IIASA Working Papers*, WP-95-061. Laxenburg, Austria: International Institute for Applied Systems Analysis.
- Nilsen, P. 2015. "Making Sense of Implementation Theories, Models and Frameworks." *Implementation Science* 10 (1): 53–65. doi:10.1186/s13012-015-0242-0.
- Nilsen, P., C. Ståhl, K. Roback, and P. Cairney. 2013. "Never the Twain Shall Meet? A Comparison of Implementation Science and Policy implementation Research." *Implementation Science* 8 (1): 63–75. doi:10.1186/1748-5908-8-63.
- O'Toole, L. J. 2000. "Research on Policy Implementation: Assessment and Prospects." *Journal of Public Administration Research and Theory* 10 (2): 263–288. doi:10.1093/oxfordjournals.jpart.a024270.
- Oates, T. 2008. "Going Round in Circles: Temporal Discontinuity as a Gross Impediment to Effective Innovation in Education and Training." *Cambridge Journal of Education* 38 (1): 105–120. doi:10.1080/03057640801890012.
- OECD. 2009. *Working Out Change: Systemic Innovation in Vocational Education and Training*. Paris, France: Organization for Economic Cooperation and Development (OECD), Centre for Educational Research and Innovation (CERI).
- OECD. 2011. *Learning for Jobs - OECD Reviews of Vocational Education and Training: Pointers for Policy Development*. Paris, France: Organization for Economic Cooperation and Development (OECD).
- Paik, S. J. (2014). "Pre-employment VET Investment Strategy in Developing Countries: Based on the Experiences of Korea." *KDI School of Public Policy and Management Working Papers* No. 14-06.
- Popper, K. (2014). "Conjectures and Refutations: The Growth of Scientific Knowledge." New York: Harper Torchbooks. (Original work published 1963).
- Powell, M. 2001. "A Comparative Study of TVET Projects—Implementation Experiences from Jamaica and the Gambia." *International Journal of Educational Development* 21 (5): 417–431. doi:10.1016/S0738-0593(00)00066-3.

- Pressman, J. L., and A. Wildavsky. 1973. *Implementation*. Los Angeles, CA: University of California Press.
- Pusterla, F. 2016. "The Great Recession and the Working Conditions of Youth: A Descriptive Analysis of the European Labour Market." *KOF Studies*, no. 83.
- Rageth, L., and U. Renold. 2019. "The Linkage between the Education and Employment Systems: Ideal Types of Vocational education and Training Programs." *Journal of Education Policy* 1–26. doi:10.1080/02680939.2019.1605541.
- Rekkor, S., M. Ümarik, and K. Loogma. 2013. "Adoption of National Curricula by Vocational Teachers in Estonia." *Journal of Vocational Education & Training* 65 (4): 489–506.
- Ryan, P. 2000. "The Institutional Requirements of Apprenticeship: Evidence from Smaller EU Countries." *International Journal of Training and Development* 4 (1): 42–65. doi:10.1111/ijtd.2000.4.issue-1.
- Schofield, J. 2001. "Time for A Revival? Public Policy Implementation: A Review of the Literature and an Agenda for Future Research." *International Journal of Management Reviews* 3 (3): 245–263. doi:10.1111/ijmr.2001.3.issue-3.
- Sohn, H., K. Kang, and J. Lim. 2017. "More Knowing, Less Doing: International Organizations' Agendas in Korean TVET Aid." *KEDI Journal of Educational Policy* 14 (1): 119–136.
- Souto-Otero, M. 2011. "Breaking the Consensus in Educational Policy Reform?." *Critical Studies in Education* 52 (1): 77–91. doi:10.1080/17508487.2011.536514.
- Stoica, A. 2003. "Vocational Education and Training Reform in Romania and Bosnia-Herzegovina: Strategy, Legislation, and Implementation." *European Journal of Education* 38 (2): 213–222. doi:10.1111/ejed.2003.38.issue-2.
- Sultana, R. G. 2008. *The Challenge of Policy Implementation: A Comparative Analysis of Vocational School Reforms in Albania, Kosovo and Turkey*. Luxembourg: Office for Official Publications of the European Communities.
- Viennet, R., and B. Pont. (2017). "Education Policy Implementation: A Literature Review and Proposed Framework." (OECD Education Working Papers No. 162). Paris, France: OECD Publishing.
- Viertel, E., and P. Grootings. 2001. *Review and Lessons Learned of Phare Vocational Education and Training Reform Programmes, 1993–1998*. Turin, Italy: European Training Foundation.
- Wallenborn, M. 2010. "Vocational Education and Training and Human Capital Development: Current Practice and Future Options." *European Journal of Education* 45 (2): 181–198. doi:10.1111/(ISSN)1465-3435.
- Wilson, D. N. 1993. "Reforming Technical and Technological Education." *The Vocational Aspect of Education* 45 (3): 265–284. doi:10.1080/0305787930450307.
- Winter, S. 1990. "Integrating Implementation Research." In *Implementation and the Policy Process. Opening up the Black Box*, edited by D. J. Palumbo and D. J. Calista, 19–38. New York, NY: Greenwood Press.
- Winter, S. C. 2012. "Implementation Perspectives: Status and Reconsideration." In *Handbook of Public Administration*, edited by B. G. Peters and J. Pierre, 265–278. London, England: Sage.
- Wolter, S. C., and P. Ryan. 2011. "Apprenticeship." In *Handbook of the Economics of Education*, edited by E. Hanushek, S. Machin, and L. Woessmann, 521–576. Vol. 3. Amsterdam, Netherlands: North Holland.
- Wolter, S. C., S. Mühlemann, and J. Schweri. 2006. "Why Some Firms Train Apprentices and Many Others Do Not." *German Economic Review* 7 (3): 249–264. doi:10.1111/geer.2006.7.issue-3.



Appendix

Table A1. Search terms and databases.

Database	Syntax	Syntax explanation	Filters	Date of search	No. of results	Coded
EBSCO Education Source	TI ('vocational education' OR VET OR TVET) AND (AB implement* OR SU implement* OR TI implement*)	Title (TI), abstract (AB), subject terms (SU)	Full text	22.1.2018	65	18
ERIC	title:(vocational education' OR VET OR TVET) AND (abstract:(change OR changes OR reforming OR reform OR reforms OR reforming OR innovate OR innovates OR innovation OR innovating) OR title:(change OR changes OR changing OR reform OR reforms OR reforming OR innovate OR innovates OR innovation OR innovating)) AND (abstract:(implementation OR implementing OR implement OR implements) OR title:(implementation OR implementing OR implement OR implements))	Publication types: reports (research/descriptive/ evaluative/general), journal articles, ERIC publications, books	Full text; Publication types: reports (research/descriptive/ evaluative/general), journal articles, ERIC publications, books	23.1.2018	185	14
VocedPlus	(tm_metadata.title:(vocational education' OR TVET OR VET) AND (tm_metadata.title:(change OR changes OR changing OR reform OR reforms OR reforming OR innovate OR innovates OR innovation OR innovating) OR tm_metadata.abstract:(change OR changes OR changing OR reform OR reforms OR reforming OR innovate OR innovates OR innovation OR innovating)) AND (tm_metadata.abstract:(implementation OR implementing OR implement OR implements) OR tm_metadata.title:(implementation OR implementing OR implement OR implements)) AND bs_metadata.fulltext:true		Full text; Resource types: reports, articles, theses, papers, working papers, books	23.1.2018	198	34
Science Direct	ttitle('vocational education' OR VET OR TVET) AND tak(implement*)	Title (ttl), title, abstract, and author or publisher's keywords (tak)		24.1.2018	27	4

(Continued)

Table A1. (Continued).

Database	Syntax	Syntax explanation	Filters	Date of search	No. of results	Coded
Web of Science	Ti=(vocational education' OR VET OR TVET) AND Ti=(implement* OR reform* OR chang* OR innovat*)	Title (Ti)	Categories: Education educational research, psychology educational, education scientific disciplines, education special, Resource types: articles, book chapters, books	25.1.2018	107	7
JSTOR	(ti:(vocational education' OR TVET OR VET) OR tb:(vocational education' OR TVET OR VET) AND implement* AND (chang* OR reform* OR innovat*))	Title (ti OR tb), Language: English		24.1.2018	132	14
SAGE	[[Publication Title 'vocational education'] OR twet] AND [[Abstract reform*] OR [Abstract chang*] OR [Abstract innovat*] OR [Abstract implement*]]		Full Access	23.1.2018	31	8
Taylor and Francis Online	[[[Publication Title: 'vocational education'] OR Title: tvet]] AND [[Publication Title: implement*] OR [Publication Title: reform*] OR [Publication Title: chang*] OR [Publication Title: innovat*]]		Full Access	25.1.2018	61	5
UNESCO-UNEVOC Publications	Title (implement; reform; innovat; chang) Wildcards (e.g. the asterik sign) are not allowed, but search for word stem (e.g. implement) also looks for variations (e.g. implementation, implementing, implements). Each term (implement, reform, innovat, chang) was searched seperately. The number of results taken together from all searches was 318.			24.1.2018	318	14

(Continued)

Table A1. (Continued).

Database	Syntax	Syntax explanation	Filters	Date of search	No. of results	Coded
OECD ilibrary	Search Results containing "vocational education" OR VET OR TVET' AND Title, Abstract, Authors, Keyword or ISSN/ISBN/ DOI containing 'implement*' Restricted to Language(s) English Restricted to Theme Education in content type BOOK SERIES OR in content type BOOK OR in content type CHAPTER OR in content type JOURNAL OR in content type ARTICLE OR in content type ANNUAL OR in content type WORKING PAPER SERIES OR in content type WORKING PAPER Published Between 1900 and 2018		Theme: Education; Content type: book series, books, chapters, journals, articles, annuals, working paper series, working papers	25.1.2018	20	13
European Commission Find-eR	title('vocational education' OR VET OR TVET) AND title('implement*' OR reform*' OR chang*' OR innovat*)			25.1.2018	140	8
ETF Publications Catalogue	- Categories ('topics'): VET System Assessment, Qualifications Systems, Learning And Teaching In VET, VET Quality Assurance, VET Governance			24.1.2018	334	19
Cedefop	- Categories ('tags'): apprenticeship, comparative analysis, EOF, education and training policy, education and training system, national policy, National Qualifications Framework, vocational education and training			24.1.2018	217	15

Table A2. Subsamples by organisation compared to overall non-peer-reviewed, development status within Europe.

Category	Organisations compared to overall non-peer-reviewed												Development Status in Europe			
	Lit – ETF	ETF	Std. Residual	Lit – Cedefop	Cedefop	Std. Residual	Lit – OECD	OECD	Std. Residual	Developed	Developing	Standard Residual	Developed	Developing	Standard Residual	
Content	48	15	0.44	58	5	-0.47	54	9	-1.60	19	27	1.61	19	27	1.61	
	39	10	-0.31	43	6	1.01	39	10	-0.41	17	13	-0.65	17	13	-0.65	
	24	5	-0.70	26	3	-0.17	24	5	-0.55	12	14	0.26	12	14	0.26	
	17	7	0.24	22	2	-0.65	20	4	-0.51	14	9	-0.94	14	9	-0.94	
	16	4	-0.47	17	3	0.30	19	1	-1.11	7	8	-0.10	7	8	-0.10	
Context	12	4	-0.23	14	2	-0.23	10	6	0.59	7	4	-0.83	7	4	-0.83	
	37	12	0.33	44	5	0.27	39	10	-0.41	13	20	1.21	13	20	1.21	
	35	10	-0.07	38	7	1.97*	43	2	-2.45**	14	13	-0.21	14	13	-0.21	
	18	9	0.70	24	3	-0.06	26	1	-1.57	7	9	0.08	7	9	0.08	
Commitment	23	12	1.18	30	5	1.01	27	8	-0.07	10	13	0.37	10	13	0.37	
	9	1	-0.83	8	2	0.09	10	0	-0.75	2	7	0.44	2	7	0.44	
	3	0	-0.72	2	1	-0.29	3	0	-0.29	1	0	-0.69	1	0	-0.69	
	37	11	0.07	41	7	1.81*	44	4	-2.07**	17	15	-0.29	17	15	-0.29	
	41	11	-0.17	47	5	0.11	43	9	-0.89	15	15	0.00	15	15	0.00	
	30	13	1.01	38	5	0.58	39	4	-1.74*	13	18	0.84	13	18	0.84	
Capacity	14	6	0.17	16	4	1.05	19	1	-1.11	10	6	-0.91	10	6	-0.91	
	7	5	0.33	10	2	-0.02	12	0	-0.88	1	6	0.41	1	6	0.41	
	46	15	0.56	56	5	-0.36	53	8	-1.76*	17	21	0.81	17	21	0.81	
	43	14	0.48	53	4	-0.90	48	9	-1.22	19	20	0.33	19	20	0.33	
	50	13	-0.20	56	7	1.02	48	15	0.12	24	20	-0.40	24	20	-0.40	
	17	2	-1.05	16	3	0.36	18	1	-1.05	15	6	-1.64*	15	6	-1.64*	
	13	6	0.23	18	1	-1.13	17	2	-0.76	4	6	-0.03	4	6	-0.03	
Clients	68	15	-0.77	78	5	-1.52	55	28	2.56**	26	23	-0.14	26	23	-0.14	
	47	12	-0.28	53	6	0.49	47	12	-0.48	19	16	-0.40	19	16	-0.40	
	58	15	-0.16	67	6	-0.25	49	24	2.06**	22	21	0.08	22	21	0.08	
	27	5	-0.88	30	2	-1.07	20	12	1.27	8	12	0.48	8	12	0.48	
	10	6	0.41	16	0	-1.72*	15	1	-0.85	1	3	-0.14	1	3	-0.14	
	27	10	0.42	33	4	0.15	34	3	-1.64*	8	14	0.85	8	14	0.85	
	21	7	0.00	25	3	-0.12	22	6	-0.19	12	11	-0.29	12	11	-0.29	
	17	7	0.24	22	2	-0.65	20	4	-0.51	7	8	-0.10	7	8	-0.10	
	854	262	R² = 0.84	1001	115	R² = 0.63	917	199	R² = 0.45	361	378	R² = 0.68	361	378	R² = 0.68	

Notes: Standardised residuals are the difference between expected and actual values divided by the standard deviation of expected values. Higher standard residuals indicate that the determinant is more common than expected in the second/right-hand/specified group.

* Significant at p<0.1, or >1.64 standard deviations from the mean ** Significant at p<0.05, or >2 standard deviations from the mean.



Table A3. Subsample by continent.

Category	Determinant	Multi-Continent						Europe						Asia						Africa					
		Total	-Multis	Multi	Residual	Total	-Europe	Europe	Residual	Total	-Asia	Asia	Residual	Total	-Africa	Africa	Residual	Total	-Africa	Africa	Residual				
		Std.		Std.		Std.		Std.		Std.		Std.		Std.		Std.		Std.		Std.					
Content	Strategy	72	-0.24	18	-0.24	44	46	0.70	81	12	0.98	81	9	0.05	9	0.05	81	9	0.05	9	0.05				
	Accountability	46	0.18	14	0.18	30	30	0.24	59	4	-0.69	57	3	-1.13	3	-1.13	59	4	-0.69	57	3	-1.13			
	Piloting	37	-0.93	6	-0.93	17	26	0.75	43	3	-0.35	39	4	-0.04	4	-0.04	43	3	-0.35	39	4	-0.04			
	Slow Pace	32	-0.33	8	-0.33	17	23	0.49	42	1	-0.95	36	4	0.09	4	0.09	42	1	-0.95	36	4	0.09			
	Bottom-Up	22	0.12	8	0.12	15	15	-0.07	32	1	-0.54	26	4	0.49	4	0.49	32	1	-0.54	26	4	0.49			
	Incremental	18	-0.45	4	-0.45	11	11	-0.15	23	2	0.15	19	3	0.42	3	0.42	23	2	0.15	19	3	0.42			
	Coordination	59	-0.22	15	-0.22	41	33	-0.22	65	12	1.64*	66	8	0.30	8	0.30	65	12	1.64*	66	8	0.30			
	Context Fit	43	1.83*	22	1.83*	38	27	-0.54	65	3	-1.25	58	7	0.27	7	0.27	65	3	-1.25	58	7	0.27			
Context	Decentralise	24	0.41	10	0.41	18	16	-0.18	34	3	0.02	30	4	0.33	4	0.33	34	3	0.02	30	4	0.33			
	Formalism	35	0.67	14	0.67	26	23	-0.10	45	7	0.85	46	3	-0.68	3	-0.68	45	7	0.85	46	3	-0.68			
	Strong Econ	12	-0.37	3	-0.37	6	9	0.01	17	1	0.08	13	2	0.30	2	0.30	17	1	0.08	13	2	0.30			
	Ed Quality	3	-0.16	2	-0.16	4	1	-0.55	7	1	0.49	4	1	0.31	1	0.31	7	1	0.49	4	1	0.31			
	Political Will	58	0.39	18	0.39	44	32	-0.50	73	6	-0.62	66	10	1.02	10	1.02	73	6	-0.62	66	10	1.02			
	Cooperation	53	0.05	15	0.05	38	30	-0.28	64	7	0.07	61	7	0.14	7	0.14	64	7	0.07	61	7	0.14			
	Foreign Aid	45	-0.35	11	-0.35	25	31	0.65	54	5	-0.16	48	8	1.03	8	1.03	54	5	-0.16	48	8	1.03			
	Ownership	24	0.22	9	0.22	17	16	-0.11	32	4	0.43	31	2	-0.43	2	-0.43	32	4	0.43	31	2	-0.43			
Capacity	Low Turnover	10	-0.09	4	-0.09	7	7	-0.23	16	1	0.12	13	1	-0.06	1	-0.06	16	1	0.12	13	1	-0.06			
	Personnel	74	0.24	21	0.24	57	38	-0.84	86	12	0.77	81	14	1.84*	14	1.84*	86	12	0.77	81	14	1.84*			
	Finances	70	-0.15	18	-0.15	49	39	-0.23	84	7	-0.75	75	13	1.73*	13	1.73*	84	7	-0.75	75	13	1.73*			
	Research	68	-0.06	18	-0.06	42	44	0.66	80	9	0.06	79	7	-0.59	7	-0.59	80	9	0.06	79	7	-0.59			
	Time	31	-0.66	6	-0.66	16	21	0.38	39	1	-0.83	32	5	0.61	5	0.61	39	1	-0.83	32	5	0.61			
	Leadership	23	-0.30	6	-0.30	19	10	-0.76	28	4	0.59	27	2	-0.27	2	-0.27	28	4	0.59	27	2	-0.27			
	Employers	85	-0.45	20	-0.45	56	49	0.17	96	12	0.36	97	8	-0.96	8	-0.96	96	12	0.36	97	8	-0.96			
	Educators	58	-0.18	15	-0.18	38	35	0.15	69	7	-0.13	68	5	-0.86	5	-0.86	69	7	-0.13	68	5	-0.86			
Clients	Intermediaries	66	0.03	18	0.03	41	43	0.64	79	8	-0.22	78	6	-0.90	6	-0.90	79	8	-0.22	78	6	-0.90			
	Trade Unions	27	0.47	11	0.47	18	20	0.17	37	4	0.22	38	0	-1.43	0	-1.43	37	4	0.22	38	0	-1.43			
	Community	11	0.24	6	0.24	13	4	-0.88	19	1	0.00	15	2	0.22	2	0.22	19	1	0.00	15	2	0.22			
	Low Level	34	0.53	13	0.53	25	22	-0.12	46	4	-0.15	44	3	-0.60	3	-0.60	46	4	-0.15	44	3	-0.60			
	High Level	30	-0.24	8	-0.24	15	23	0.62	38	3	-0.14	36	2	-0.63	2	-0.63	38	3	-0.14	36	2	-0.63			
	Mid Level	21	-0.21	6	-0.21	12	15	0.13	28	2	-0.05	26	1	-0.58	1	-0.58	28	2	-0.05	26	1	-0.58			
	Totals/R ²	1191		347	R² = 0.87	799	739	R² = 0.94	157	147	R² = 0.78	159	148	R² = 0.70	148	R² = 0.70	159	148	R² = 0.78	159	148	R² = 0.70			

Notes: Standardised residuals are the difference between expected and actual values divided by the standard deviation of expected values. Higher standard residuals indicate that the determinant is more common than expected in the second/right-hand/specified group.

* Significant at $p < 0.1$, or > 1.64 standard deviations from the mean ** Significant at $p < 0.05$, or > 2 standard deviations from the mean.

Table A4. Covariance matrix.

Content	Content						Context						Commitment				
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
Content	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5
	0.25	0.22															
	0.03	0.02	0.18														
	0.01	0.02	0.04	0.17													
	0.02	0.01	0.02	0.02	0.14												
	0.04	0.00	0.02	0.04	0.03	0.11											
	0.00	0.01	0.02	0.04	0.01	0.01	0.24										
Context	0.08	0.04	0.03	0.04	0.03	0.04	0.10	0.23									
	0.04	0.02	0.04	0.02	0.05	-0.01	0.04	0.06	0.16								
	0.03	0.03	0.01	0.01	0.04	0.03	0.10	0.03	0.01	0.20							
	0.05	0.05	0.05	0.02	0.01	0.01	0.03	0.03	0.01	0.02	0.08						
	0.01	0.01	0.00	0.00	0.01	0.01	0.03	0.03	0.01	0.01	0.01	0.03					
	0.01	0.01	0.00	0.00	0.02	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.03				
	0.01	0.00	0.00	0.05	0.04	0.01	0.03	0.06	0.03	0.03	0.02	0.02	0.01	0.03			
Commitment	0.08	0.02	0.02	0.02	0.04	0.00	0.06	0.04	0.04	0.05	0.05	0.05	0.25	0.24	0.22	0.15	0.07
	0.05	0.08	0.03	0.02	0.00	0.01	0.05	0.08	0.05	0.05	0.05	0.02	0.04	0.02	0.02	0.00	0.02
	0.04	0.00	0.03	0.02	0.00	0.01	0.04	0.04	0.00	0.02	0.02	0.00	0.06	0.04	0.04	0.03	0.00
	0.04	0.02	-0.01	0.01	0.02	0.00	0.04	0.04	0.00	0.02	0.01	0.00	0.02	0.01	0.03	0.04	0.02
	0.02	0.02	0.03	0.02	0.02	0.01	0.02	0.03	0.02	0.02	0.02	0.00	0.02	0.01	0.04	0.04	0.03
Capacity	0.06	0.02	0.02	0.02	0.03	0.02	0.06	0.06	0.04	0.03	0.02	0.00	0.07	0.03	0.04	0.04	0.03
	0.06	0.04	0.03	0.03	0.05	0.02	0.10	0.05	0.03	0.07	0.02	0.01	0.06	0.05	0.05	0.03	0.02
	0.05	0.06	0.03	0.01	0.01	-0.02	0.06	0.06	0.06	0.05	0.02	0.00	0.05	0.04	0.05	0.03	0.02
	0.03	-0.02	0.02	0.03	0.02	-0.01	0.01	0.01	0.00	-0.01	-0.01	0.00	0.02	0.00	-0.01	0.01	0.00
	0.02	0.02	0.02	0.02	0.02	0.00	0.03	0.03	0.01	0.01	0.00	0.00	0.03	0.04	0.03	0.04	0.01
	0.03	0.05	0.02	0.00	0.00	0.01	0.04	0.01	0.03	0.04	0.01	0.01	0.04	0.06	0.00	-0.01	0.02
Clients	0.03	0.05	0.03	0.02	0.00	0.00	0.05	0.03	0.03	0.03	0.00	0.00	0.03	0.04	0.00	0.01	0.02
	0.03	0.05	0.03	0.02	0.01	0.00	0.05	0.03	0.03	0.03	0.00	0.00	0.02	0.04	0.02	0.01	0.02
	0.04	0.06	0.03	-0.01	0.01	0.00	0.06	0.04	0.04	0.05	0.00	0.00	0.02	0.07	0.04	0.01	0.02
	0.04	0.05	0.02	0.00	0.01	0.00	0.02	0.02	0.05	0.03	0.01	0.01	0.02	0.04	-0.01	0.00	0.01
	0.02	0.02	0.00	0.01	0.02	0.01	0.02	0.03	0.02	0.00	0.01	0.00	0.02	0.03	0.02	0.00	0.02
	0.05	0.05	0.03	0.01	0.04	0.02	0.05	0.06	0.05	0.02	0.01	0.00	0.04	0.06	0.03	0.05	0.02
	0.02	0.03	0.04	0.01	0.01	0.00	0.01	0.03	0.02	0.02	0.00	0.00	0.03	0.04	0.02	0.01	0.02
	0.02	0.03	0.04	0.01	0.04	0.02	0.01	0.03	0.02	0.03	0.00	0.00	0.03	0.04	0.03	0.05	0.02
	0.04	0.04	0.03	0.01	0.02	-0.01	0.04	0.03	0.04	0.02	0.00	0.00	0.01	0.05	0.02	0.01	0.02
	0.04	0.04	0.03	0.01	0.02	0.00	0.04	0.03	0.04	0.00	0.00	0.00	0.01	0.05	0.03	0.02	0.02
Average	0.04	0.03	0.02	0.02	0.02	0.01	0.04	0.04	0.03	0.03	0.01	0.01	0.03	0.04	0.03	0.02	0.02
Dimension Average	0.028					0.031							0.032				

(Continued)



Table A4. (Continued).

	Capacity					Clients								
	1	2	3	4	5	1	2	3	4	5	6	7	8	
Content	1													
	2													
	3													
	4													
	5													
Context	1													
	2													
	3													
	4													
	5													
Commitment	1													
	2													
	3													
	4													
	5													
Capacity	1	0.25												
	2	0.12	0.25											
	3	0.04	0.06	0.25										
	4	0.02	0.02	0.00	0.17									
	5	0.03	0.03	0.03	0.01	0.14								
Clients	1	0.00	0.04	0.06	0.06	0.02	0.24							
	2	0.05	0.04	0.06	-0.01	0.05	0.11	0.24						
	3	0.02	0.03	0.07	-0.05	0.01	0.16	0.12	0.25					
	4	0.00	0.01	0.05	-0.03	0.00	0.09	0.04	0.10	0.17				
	5	0.04	0.04	0.02	0.01	0.03	0.03	0.04	0.03	0.00	0.09			
	6	0.07	0.07	0.07	-0.02	0.04	0.02	0.06	0.04	0.02	0.04	0.20		
	7	0.01	0.03	0.03	-0.01	0.02	0.06	0.05	0.06	0.04	0.03	0.07	0.17	
	8	0.02	0.03	0.04	-0.01	0.03	0.03	0.02	0.04	0.04	0.02	0.08	0.13	
Average		0.03	0.04	0.04	0.01	0.02	0.04	0.04	0.04	0.03	0.02	0.04	0.03	0.03
Dimension Average		0.037					0.034							

Notes: Averages are for the absolute value of covariances. The values along the diagonal are within-determinant variances, and values filling the matrix are covariances between determinants. Highlighted cells are above 0.075 (yellow) or 0.1 (red).