Iannacci, F. and Comford, T. (2017) Unravelling causal and temporal influences underpinning monitoring systems success: a typological approach. Information Systems Journal. ISSN 1350-1917.

http://dx.doi.org/10.1111/isj.12145

UNRAVELLING CAUSAL AND TEMPORAL INFLUENCES UNDERPINNING MONITORING SYSTEMS SUCCESS: A TYPOLOGICAL APPROACH

ABSTRACT

This paper is concerned with the causal and temporal underpinnings of Information Systems (IS) success. **It uses a typological approach** based on fuzzy-set Qualitative Comparative Analysis (fsQCA) and process tracing. It investigates success **across** multiple cases of IS adopted for monitoring the disbursement and use of resources **within** the European Social Fund (ESF). The study unravels the causal mechanisms and temporal pathways underpinning success in these systems. **It develops a typological theory of monitoring systems success** that reveals the temporal pathways embedded within individual cases, as well as broader theoretical patterns emerging across cases. Theoretical, methodological and practical implications are discussed.

Key words: IS success, monitoring systems, qualitative comparative analysis, process tracing, typological theorising.

1. INTRODUCTION

The study of information systems (IS) success is a recurrent issue that raises **both** theoretical and methodological challenges (DeLone and McLean, 1992; 2003; Dwivedi et al., 2015). Theoretically, these include overcoming simplistic assumptions of linear relationships. Methodologically, it is difficult to unpack the causal processes that link independent variables (or causal conditions) with the outcome of interest in particular contexts (Caldeira and Ward, 2002; Henfridsson and Bygstad, 2013). **We take on these challenges here using a typological approach that integrates fuzzy-set Qualitative Comparative Analysis (fsQCA) with process tracing in a theoretically-pluralist fashion.**

Our typological approach contributes several powerful ideas to the IS success literature. First, the relationships between causal conditions and an outcome of interest need not be treated as crudely linear (e.g., "if more X then more Y". Cf. Markus and Robey, 1988: 590). On the contrary, causal conditions work together rather than separately because they are parts of more complex systems that produce holistic effects that may not be inherent in their individual parts (Mahoney, 2001). Second, the combinations (or configurations) of causal conditions leading to monitoring systems success need not be simply the reverse of the configurations for failure. This, in turn, challenges the common correlational assumption of causal symmetry (Fiss, 2011; Liu et al., 2017). Third, there can be multiple, equally-effective pathways to IS success (i.e., equifinality). These pathways can encompass a variety of types ranging from a straightforward chain of events to a nexus of co-occurring events. By identifying the causal conditions mobilised in these pathways, IS success scholars can develop typological theories, that is, contingent generalisations about configurations of conditions that constitute theoretical types (George and Bennett, 2005).

The approach we advocate **here** offers IS success researchers the opportunity to study a diverse and even eclectic range of IS and discover therein different types, subtypes and mixed types of configurations (Ragin, 2000). Rather than focusing IS success research on the 'net effects' of causal conditions working independently of each other as 'variables', our approach offers a configurational view **premised on the assumption that causal conditions are embedded within more complex systems (Burton-Jones et al., 2015; El Sawy et al., 2010; Mingers, 2014).** Accordingly, researchers can examine the multiple, logically-possible ways that causal conditions may combine to produce an outcome of interest and express this as a theory of types (Fiss, 2011).

There are indeed many situations where identifying equifinal configurations of multiple conditions affecting a specific outcome can offer real insights **regarding** the interdependence of conditions and their holistic effects. From there, a theorisation of the multiple ways that an outcome might be achieved is possible. Beyond this, our novel approach provides insight into the processes (including **chains of events** or co-occurring events) that mobilise **distinct mechanisms** and underpin equifinality. This understanding can serve the practical interests of those who manage or commission the evaluation of IS (Klecun et al., 2014). It can also serve those who study classes of IS in the field as a basis for typological theorising (Avgerou, 2013; George and Bennett, 2005). Typological theorising is a valuable if underappreciated **approach** in IS success research, one that can shift attention to theories of the middle range that capture the intricacies of specific processes, as well as the general theoretical patterns evident across cases (Ibid).

Methodologically, our study integrates fsQCA with process tracing within a sequential, multi-method research design (Crilly, 2011; George and Bennett, 2005). The cases we compare are **of** monitoring systems established and run by European governments in the course of their implementation of the European Social Fund (ESF) - a major European Union (EU) socio-economic support programme. Each case constitutes country-wide, or at times, region-wide, arrangements for monitoring processes and data flows spanning multiple organisations. Drawing from common EU regulations, each country or region must enact some standards for the required data gathering and undertake some specific validation tasks. Overall, these are complex multi-organisational systems that embody technical, legislative, organisational and social features. Assessing their success is not easy. Our research question, presented in two parts, is then:

> 1) What aspects of these monitoring systems are relevant for a positive effect (i.e., outcome, consequence or impact) to be seen? 2) **In what ways do** such aspects produce these effects?

To address this two-part question, we selected comparable cases within the EU. We used these cases to develop a nuanced explanation of monitoring systems success that is responsive to novel methodological developments in the IS field (El Sawy et al., 2010; Henfridsson and Bygstad, 2013; Liu et al., 2017; Park and El Sawy, 2013; Rivard and Lapointe, 2012; **Tan et al., 2016**). To this end, we re-interpreted the DeLone and McLean (1992) model of IS success with the structure-process-outcome framework proposed by Cornford et al. (1994) and unravelled the causal and temporal influences underpinnings monitoring systems success in a contingent fashion.

The remainder of the paper unfolds as follows. Section two introduces **Cornford's et al. (1994) framework** and provides the theoretical and methodological backdrop against which the original and the updated versions of the DeLone and McLean model are reviewed and fsQCA is deployed in combination with process tracing. Section three and four sketch our coding and calibration procedures with regard to our causal conditions, i.e., information and system quality components, and outcome, i.e., impact, respectively. Section five presents a **typological theory of monitoring systems success that reveals the temporal processes embedded within individual cases, as well as the causal configurations** for positive and negative impacts at both an abstract and more granular level of analysis. Section six summarises our work and offers concluding remarks on the theoretical, methodological and practical contributions of this study, as well as its limitations. **Online appendices give a list of key sources and activities, as well as a summary of the core features of the cases under investigation and details regarding the calibration procedures.**

2. THEORETICAL AND METHODOLOGICAL BACKGROUND

This research was inspired by participation in an evaluation project to assess the relative success of the monitoring systems established by European governments at national or regional levels for ESF activities in the 2000-2006 programming period (references withheld). The ESF is the principal European Union (EU) policy instrument to achieve economic cohesion. It operates through a set of programmes in all EU countries which are funded by the European Commission. It had a budget of approximately 60 billion Euros over the period 2000-2006. The ESF was then and continues to be focused on a consistent set of objectives (e.g., matching labour market demand and supply, creating net jobs, etc.) and funds corresponding projects (e.g., training projects). As part of the implementation of this policy, **all** ESF-funded projects in each EU country are subjected to monitoring to ascertain the veracity of claims for funding with regard to what is being achieved and how much money is being invested. This paper is concerned with assessing the relative success of these monitoring systems, each of which is different in many ways **mirroring** the various national or regional contexts, though serving the same ends and reflecting the single EU-wide regulation, Regulation 1260/99.

These monitoring systems support inspection and audit of ESF-funded projects including their inputs, outputs, results and impacts. For example, for a project providing training (the most common type), data is required on the amount of financial resources used (i.e., inputs), the number of training hours delivered (i.e., outputs), the number of successful trainees (i.e., results) and the number of unemployed people finding stable employment (i.e., long-term impacts). Based on the common schema (i.e., Regulation 1260/99), ESF-funded projects are monitored by means of physical and financial indicators. The former refer to synthetic summary metrics tracking project implementation (e.g., number of training places, number of training hours, number of successful trainees, etc.). The latter refer to financial resources used to implement projects (e.g., financial inputs, eligible costs, etc.). This information is aggregated at national levels and delivered to the European Commission. Figure 1 depicts the different layers of responsibility for monitoring of ESF-funded projects.

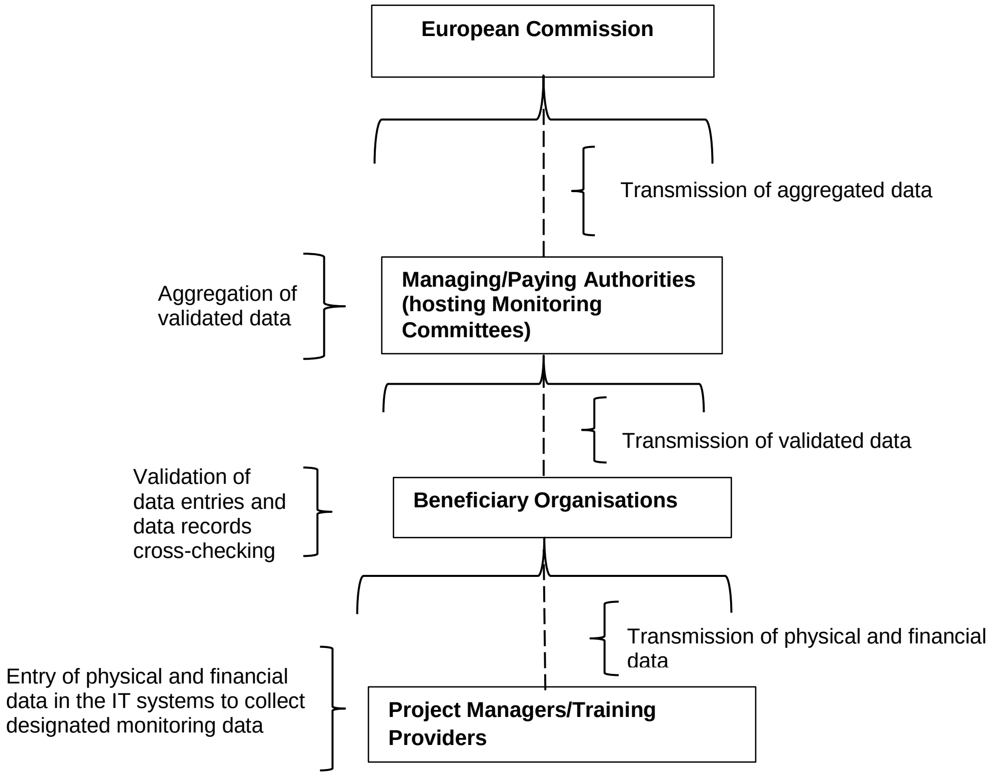

Figure 1: Archetypical flow of ESF monitoring data. Adapted from the Practical Guidance on Data Collection and Validation issued by the European Commission: Employment, Social Affairs & Inclusion (2015).

Officially, responsibility for oversight and auditing belongs to the European Commission (i.e., the Regulator) and the Monitoring Committees located within National and/or Regional Managing Authorities. These organisations, in turn, **pass** the oversight-related data gathering and data transmission workload to Beneficiary Organisations who then **pass** it to Project Managers (or Training Providers). This complex web of relationships exists within the context of Regulation 1260/99 that defines standards for data gathering and legitimate practices of communication and exchange of monitoring data between and among the various stakeholders. Accordingly, a country's or region's monitoring system is a complex and distributed socio-technical system set within its public administration (Lamb and Kling, 2003; Lee, 2010; Mumford, 2006). The purpose it serves is collecting monitoring data, transmitting and validating it. Each monitoring system embeds specific work processes and is itself embedded within wider institutional and technological structures. Each system produces positive and negative consequences and operates with varying degrees of efficiency depending on whether or not monitoring data has been appropriately collected, validated and transmitted.

Given this research context, we regard Cornford's et al. (1994) structure-process-outcome framework as very fitting. This framework sees the sequential flow of information as being structured by pre-existing contextual conditions that are mobilised when various stakeholders engage in social interactions (or processes). When stakeholders interact to collect monitoring data, transmit and validate it, they draw on pre-existing structural conditions and produce positive or negative downstream consequences (or effects, outcomes or impacts) depending on whether or not the underlying process is enacted in an efficient and effective fashion. Figure 2 introduces Cornford's et al (1994) framework.

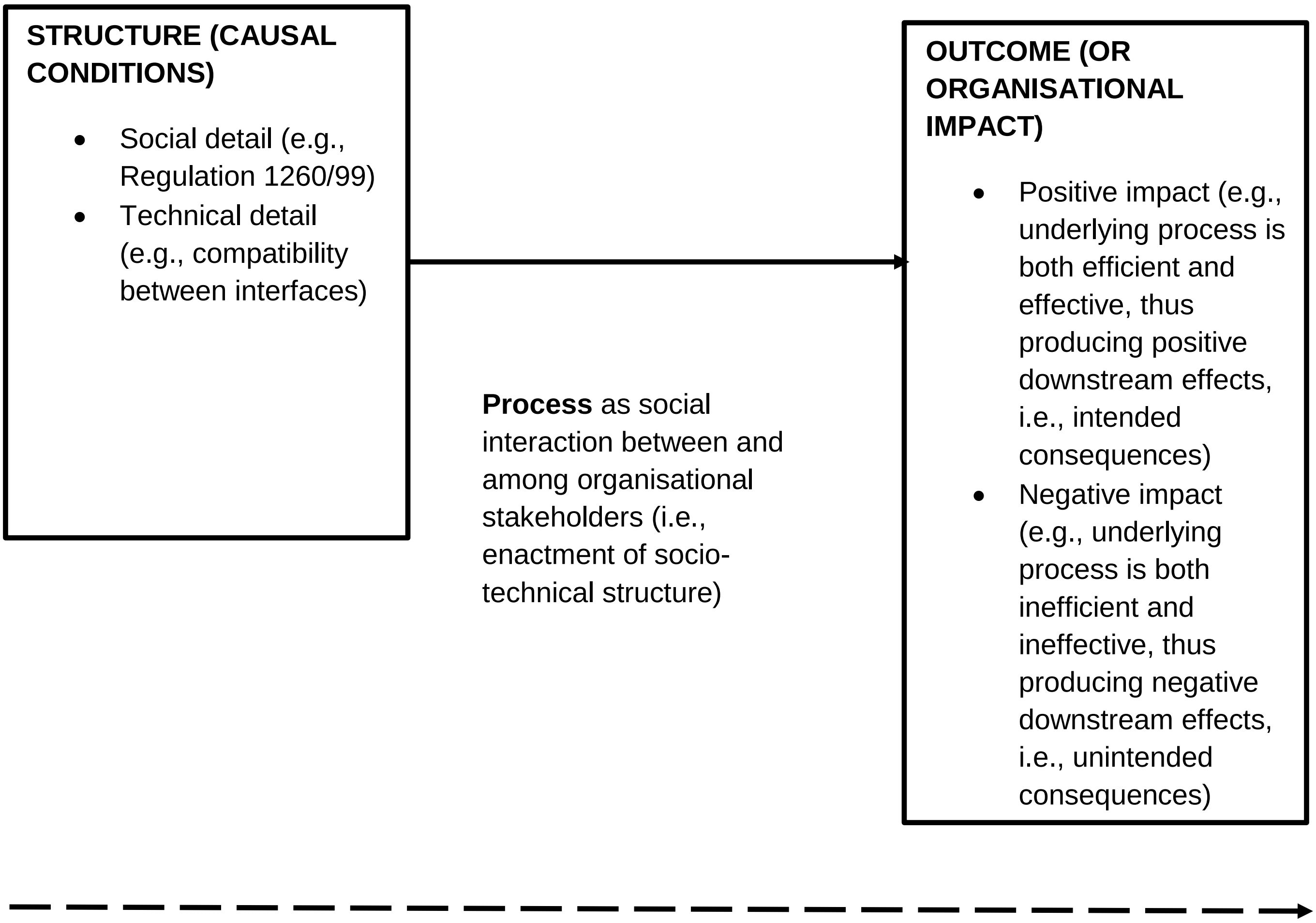

Figure 2: The structure-process-outcome framework (Source: Cornford et al., 1994).

In the following sections, we show how Cornford's et al. (1994) framework used for data collection and within-case analysis complements the DeLone and McLean model of IS success. In particular, Cornford's et al. (1994) framework helps us to untangle the causal link from the arrow of time in a configurational analysis that approaches causality from a systemic perspective (Burton-Jones et al., 2015; El Sawy et al., 2010; Mingers, 2014).

2.1. Comparing the original and updated IS success model within the ESF context

Though not without its critics (Ballantine et al., 1996; Seddon 1997), the DeLone and McLean model has become a pillar in IS success research (**DeLone and McLean, 2016**; Petter et al., 2008; 2012; 2013; Urbach et al., 2009). Informed by theories of communication (Shannon and Weaver, 1963) and information influence (Mason, 1978), the DeLone and McLean model argues for the multi-dimensional and interdependent nature of the IS success construct. According to this model, IS success is a dependent variable that includes both temporal and causal influences. IS success derives from the interaction of six interdependent dimensions where the intrinsic quality of the system and its information output influence the use of the system and perceptions of user satisfaction which, in turn, have a downstream impact on individual and organisational performance (DeLone and McLean, 1992: 83-87). Figure 3 depicts the original IS success model.

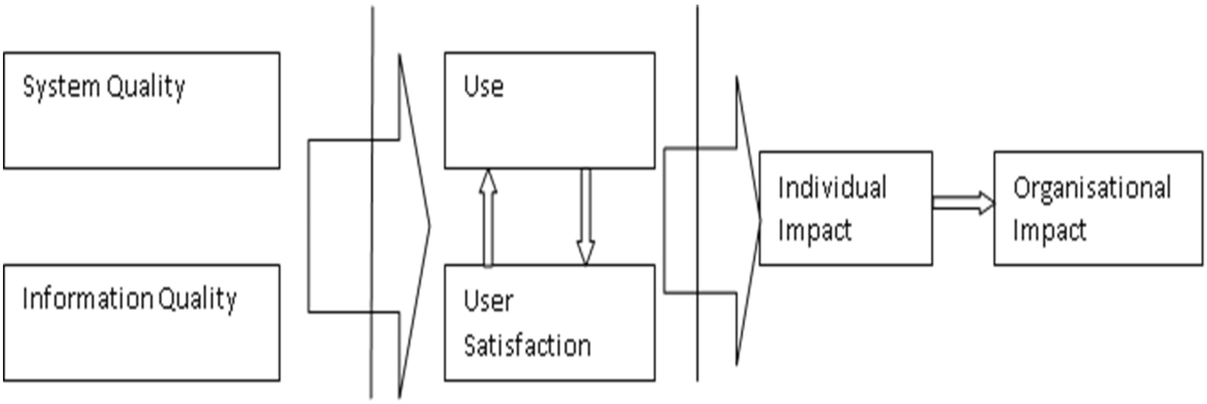

Figure 3: The DeLone and McLean model of IS success (Source: DeLone and McLean, 1992).

Ten years on the authors updated this model with the addition of 'service quality' and the collapsing of individual and organisational impacts into 'net benefits' (DeLone and McLean, 2003). Service quality was introduced to capture the overall support delivered by the IS department, and net benefits refer to downstream benefits from the perspective of the owner or sponsor of the information system. DeLone and McLean also incorporated 'intention to use' as a proxy for attitudes towards systems use and argued that "use must precede user satisfaction in a process sense, but positive experience with use will lead to greater user satisfaction in a causal sense. Similarly, increased user satisfaction will lead to increased intention to use, and thus use" (Ibid: 23). With both the original and the revised versions of the model available, we choose the original model (DeLone and McLean, 1992), rather than the updated version (DeLone and McLean, 2003). This is for three principal reasons.

First, the updated version in adding a third construct of 'service quality' seems to complicate the original model, a view **acknowledged** by the authors themselves when they suggest that service quality could be seen as "merely a subset of the model's system quality" (DeLone and McLean, 2003: 18). Others have noted that there is little conceptual gain from the updated model when the focus is "on the system as opposed to the IT function" delivered by the IS department (CF. Gable et al., 2008: 383/386; Rosemann and Vessey, 2008: 19).

Second, while the 'intention to use' construct may be a worthwhile measure in some contexts, it may be less worthwhile in other contexts (Cf. DeLone and McLean, 2003: 23). In our context, users see themselves as professionals performing roles in different organisations (e.g., the role of Project Manager providing training sessions, the role of Beneficiary Organisation granting financial resources to Project Managers by commissioning specific training projects, the role of Monitoring Committee keeping an oversight on financial resources used and targets achieved, etc.). We see these users as social actors and focus on their interactions rather than the intentions of single individuals (Lamb and Kling, 2003). Moreover, within these complex social settings, the use of electronic systems for the purpose of verifying claims for funding is mandatory. Hence, the 'intention to use' construct is not a worthwhile measure in our research context because these social actors must follow specific socio-technical protocols regardless of their own individual preferences.

Third, the idea of 'net benefits' is a catch-all concept that needs to be qualified in a contingent fashion (DeLone and McLean, 2003) and captures the ultimate impact of the information system (DeLone and McLean, 2004). Given our focus on social actors and their interactions and given the mandatory nature of the monitoring system, we choose to be more specific. For our situation, the benefit stemming from the use of the monitoring system needs to be calibrated in a way that reflects the perceived value of the monitoring system for the organisation(s) in charge (i.e., the Managing/Paying Authority and its associated Monitoring Committee). Accordingly, we assess the degree to which the monitoring system is valuable for relevant stakeholders both in terms of efficiency (i.e., tangible cost savings stemming from faster and more accurate information flows) and effectiveness (i.e., satisfied stakeholders) (Cf. Scott et al., 2016: 18; Smithson and Hirschheim, 1998: 165-166).

2.2. Research strategy and data analysis technique

Our research strategy **integrates** fuzzy-set Qualitative Comparative Analysis (fsQCA) with process tracing to unravel "the causal and temporal influences in determining IS success" (DeLone and McLean, 1992: 83). We operationalised our approach using a research design interweaving within-case and cross-case analysis in a sequential fashion (Crilly, 2011; George and Bennett, 2005).

In the first stage, **drawing on Cornford's et al. (1994) structure-process-outcome framework,** case experts gathered primary data through semi-structured interviews conducted with purposefully-selected informants (see Appendix 1). During this stage, data collection and analysis took place in parallel leading to the development of an in-depth case study for each country/region with emerging codes that explained why a monitoring system was deemed as successful or not (**see Appendix 2**). This within-case analysis was subsequently corroborated with backward coding to re-examine previous events for overlooked relevance and distil patterned similarities and differences across cases. Again new codes were generated which sometimes required the gathering of new data (mostly secondary) to validate them. Through an iterative dialogue between our fledging theoretical ideas and the empirical data, we derived a number of plausible causal conditions and outcomes in a more inductive fashion as discussed below (see Sections three and four). For each causal condition and for the chosen outcome, we developed a 'theoretical ideal' (e.g., a hypothetical case that captures all types of indicators at source, a case that uses fully-reliable monitoring systems, etc.). We used these 'theoretical ideals' as a yardstick for calibrating data from **our** actual cases (i.e., assigning fuzzy-set membership scores to all causal conditions and the chosen outcome).

In the second stage, we moved from within-case analysis to a cross-case analysis using fsQCA being mindful that process tracing can add leverage for thick cross-case comparisons (Cress and Snow, 2000; De Meur et al., 2009; Rihoux and Lobe, 2009). In this way, we identified commonalities within the same types of cases and differences across distinct types of cases (Ragin, 1987). Compared with its crisp-set variant (which uses binary data), the fuzzy-set technique allows for degrees of set membership to be specified and thus can capture more nuanced causal relationships. For example, rather than dichotomising membership in the set of cases with positive impact (i.e., presence of positive effect or absence), it allows for different membership scores between 0 and 1, thus permitting partial membership in the target set. During this second stage, analysis proceeded in four steps (Rihoux and Lobe, 2009; Schneider and Wagemann, 2012).

Step 1: Calibrating data and converting the distribution matrix into **a** truth table.

After causal conditions and the outcome of interest were calibrated (Tables 1, 3 & Appendices 3, 4 & 5), the calibration scores were aggregated (Table 4) and used to construct a distribution matrix (Table 5). This matrix displays each country's/region's fuzzy-set membership score for all combinations of aggregated conditions. Then, each country/region was assigned to that combination of aggregated conditions in which its membership score exceeded 0.50. With this data, we created a truth table to display all logically-possible combinations of aggregated conditions whether populated with cases or not (Table 6).

Step 2: Determining the outcome value for each truth-table row.

In the second step, we derived the outcome value for each truth-table row. This was based first on setting a minimum number-of-cases threshold for each row. Having a small number of cases, we set the minimum number of required cases in a row at 1. We then considered the consistency level for each row populated with cases. Consistency (sufficiency) is a measure that gauges the degree to which a causal condition or combination of conditions is a subset of the outcome. We set the lowest level of acceptable consistency at 0.85 as recommended by Ragin (Cf. Ragin, 2008: 136). Accordingly, truth-table rows with consistency above 0.85 were given a score of 1 since the configuration was a consistent subset of the outcome, or 0 if not. The empty row in the truth table was labelled a 'remainder' because it did not meet our minimum number-of-cases threshold.

Step 3: Minimising the truth table.

In the third step, we used counterfactual analysis to minimise the truth table. **Based on 'what if' claims about the outcome of the remainder**, we arrived at two solutions, **specifically**, two statements of logical conditions leading to the outcome of interest:

1) A complex solution where no 'remainders' were included and which only drew on empirical data. However, based on substantive and theoretical knowledge we moved beyond the constraints of this solution (see Section five);

2) A more parsimonious solution that included the empty truth-table row (i.e., the 'remainder' or 'empty configuration') on the assumption that it exhibited the outcome of interest. This counterfactual, in turn, permitted the elimination of redundant conditions and, therefore, the generation of a simplified solution based on the remaining **aggregate** conditions only (see Tables 7 & 8).

Step 4: Interpreting findings and process tracing.

Last, and reflecting the balance in fsQCA between cross-case and within-case analysis, we returned to the individual cases to trace the temporal pathways leading

to the outcome of interest (Cress and Snow, 2000; De Meur et al., 2009; Rihoux and Lobe, 2009). By pinpointing the causal conditions mobilised in these temporal pathways, we were able to discover the causal mechanisms underpinning monitoring systems success and conceptualise these mechanisms as complex interactions between and among causal conditions producing holistic effects not inherent in any one of them (Mahoney, 2001). We used the simplified solution terms as scope conditions for unravelling these temporal pathways and uncovering more granular causal configurations indicating different types of successful monitoring systems or the lack thereof (see Tables 9 & 10). All findings were finally validated with the fsQCA 2.5 programme.

3. CALIBRATING INFORMATION AND SYSTEM QUALITY COMPONENTS

The qualitative coding procedures used were informed by the original DeLone and McLean model of IS success interwoven with our structure-process-outcome framing for empirical research (Cornford et al., 1994). Based on this blend of theories, we identified broadly-interrelated themes by assuming that information quality and system quality contributed to positive impact either jointly or separately. We subsequently developed more nuanced conceptualisations through a back-and-forth cycling between theoretical ideas and empirical evidence (Braun and Clarke, 2006; Ragin, 2000; Rihoux and Lobe, 2009), thus progressively refining theoretical and substantive knowledge and generating in our coding more fine-grained components of information and system quality.

The resulting concept of information quality captured the various facets of the indicators generated by the systems (i.e., comprehensiveness, consistency and currency), while the system quality concept represented the more technical character of the monitoring system (i.e., compatibility, reliability and automation). Finally, the concept of impact represented the consequences (or **outcomes**) stemming from the use of the monitoring system conceived as a socio-technical system. Figure 4 provides a graphical representation of this causal model.

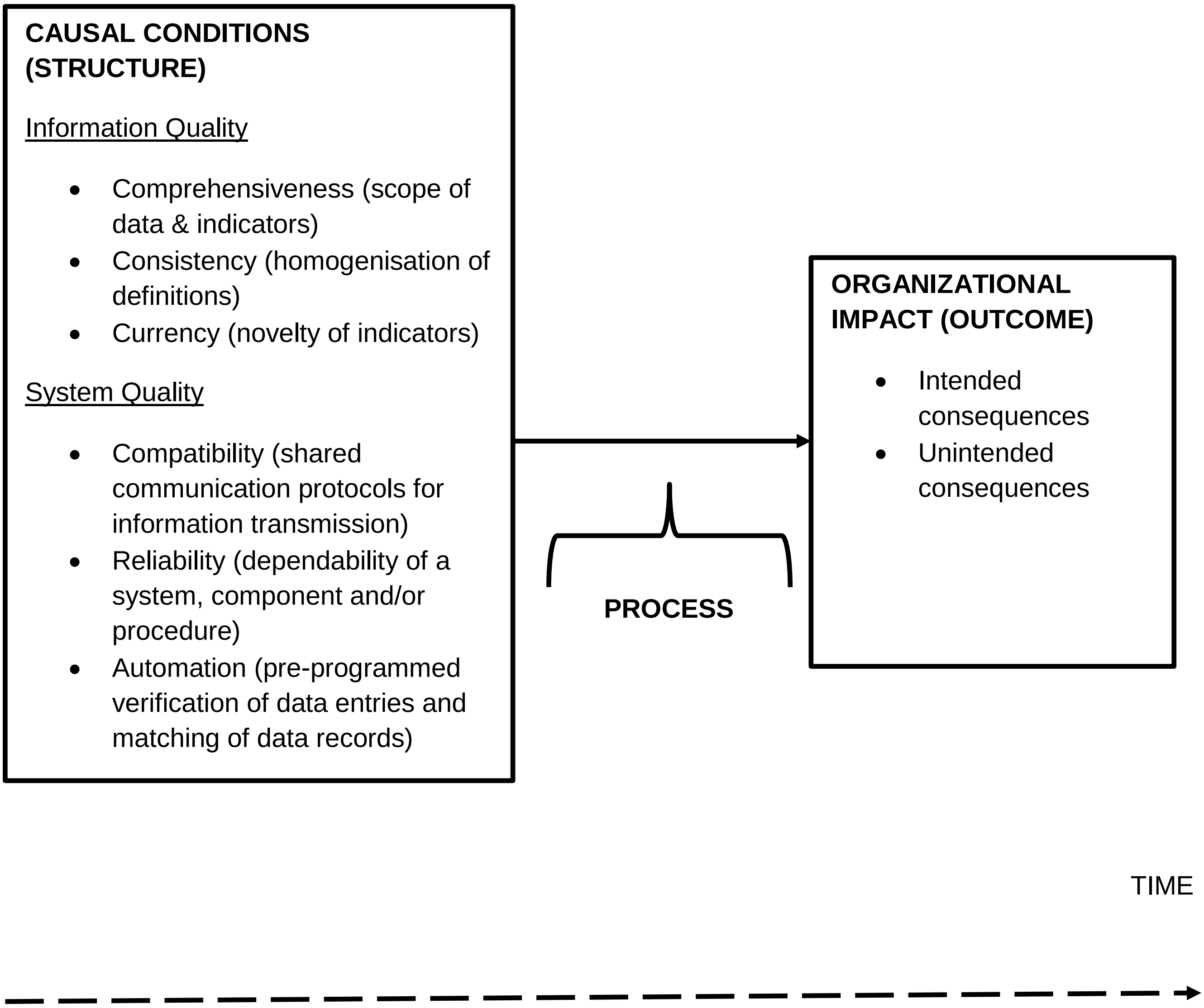

Figure 4: Graphical representation of the causal model. Adapted from Cornford et al. (1994) and DeLone and McLean (1992).

Subsequently, for each information or system quality component and for our outcome, i.e., impact, we inductively derived a 'theoretical ideal' as "the best imaginable case in the context of the study that is logically and socially possible" (Basurto and Speer, 2012: 166). We used these 'ideal types' as external standards against which our empirical cases could be calibrated (see Appendix 3). Based on individual summary statements for each case (Basurto and Speers, 2012) and Ragin's (2000: 168) advice to use five-value schemes when data is too weak to support fine-grained distinctions, **we ranked the cases under investigation (i.e., Austria, England, Flanders, France, Germany, Greece and Hungary)** in the following fashion:

- Cases that fully share the inductively-derived criteria of the ideal type, scored 1. For example, Austria scores 1 with regard to the consistency and currency of its indicators because of its concerted approach to monitoring which warrants consistently-defined input, output and result indicators, as well as indicators that are systematically updated in accordance with new information needs. Likewise, Germany scores 1 with regard to the reliability of its monitoring system because it is fully functional and there are virtually no data losses and no systems breakdowns;

- Cases that meet the inductively-derived criteria of the ideal type in a substantive fashion, scored 0.75. For example, Germany scores 0.75 with regard to the automation of its validation procedures because its monitoring system is endowed with a built-in plausibility check that examines data inputs with a colour code of green, yellow or red depending on their quality. Nevertheless, there are also manual cross-checks aimed at verifying data entries and ascertaining that financial claims match actual costs;
- Cases that partially meet some of the inductively-derived criteria of the ideal type while lacking other crucial criteria, scored 0.49[1]. For example, England scores 0.49 with regard to the comprehensiveness of its indicators because our informants acknowledged difficulties to retrieve the number of training hours despite claims that they are available in the 'wider system'. Likewise, Austria scores 0.49 with regard to the automation of its validation procedures because of its conspicuous use of manual validation routines (i.e., 'four-eye' principle) within Beneficiary Organisations despite the ongoing implementation of **a** SAP module for automated monitoring purposes. Similarly, Flanders scores 0.49 with regard to the compatibility of its monitoring system because of the lack of structured data concerning the exchange of information between the Client Following System (CFS) and the Microsoft Suite. Though data cannot be seamlessly transmitted between these two systems, the Managing Authority's Partner Institutions can still exchange email attachments by means of standard digital protocols. Furthermore, there is a pronounced use of manual validations (e.g., quality control reports) despite automated validations enabled by the CFS, thus justifying a 0.49 score with regard to the degree of automation of the monitoring system. By the same token, Hungary scores 0.49 in terms of compatibility because, despite shared digital protocols, the development of

[1] Ragin (2008: 131; footnote 2) recommends avoiding the use of 0.50 because "any case with a code of 0.50 on a causal condition will not be closest to any single corner of the vector (or property/truth-table) space defined by the causal conditions". However, scholars have bypassed this technical issue with a 0.01 adjustment (Cf. Crilly et al., 2012: 1438; Fiss 2011: 407). Accordingly, we have used a -0.01 adjustment based on the re-examination of two cases (i.e., Flanders and Hungary) that displayed different outcome values despite sharing the same values on their aggregated conditions (see Table 4). This adjustment was triggered by the bottlenecks stemming from the transmission of monitoring data in Flanders.

new modules in the Unified System has hindered the automated extraction and transmission of data;

- Cases that only meet a minority of the inductively-derived criteria of the ideal type, scored 0.25. For example, France scores 0.25 in terms of reliability because the Managing Authority's system (i.e., Application FSE) suffers from frequent breakdowns that are exacerbated by softer issues concerning the certification of expenses. Likewise, Greece scores 0.25 on this component. Not only does the Integrated Data Warehousing System suffer from legacy issues and repeated failures. It does not perform summations of physical data and decimal inputting, thus triggering manual aggregations of physical data;
- None of the cases was scored 0 because none of them was fully out of the target sets being investigated. For example, though the consistency of indicators was a challenging requirement across the board, all cases have already embarked on an effort to remove inconsistent definitions.

Table 1 shows the membership scores for each country or region with regard to each component of information and system quality. **Appendix 4 provides the summary statements associated with these scores and Appendix 5 presents** the intrinsic features of the information and system quality components with exemplary codes.

Country/ Region	**INFORMATION QUALITY**			**SYSTEM QUALITY**		
	Comprehensiveness (Ideal Type: No indicators missing; i.e., financial, output, result and impact indicators are present)	**Consistency** (Ideal type: All indicators have consistent definitions. The system of indicators is based on a concerted approach to monitoring as set out by the EU; i.e., consistently-defined input, output and result indicators)	**Currency** (Ideal Type: All indicators are regularly collected and updated; i.e., all indicators are recorded in a regular fashion and updated in accordance with new information needs)	**Compatibility** (Ideal Type: IT systems are fully compatible; i.e., able to communicate thanks to transmissions of structured data, well laid out data standards and interoperability across interfaces)	**Reliability** (Ideal Type: IT systems, components and/or procedures are fully dependable; e.g., no data losses, no systems breakdowns, seamless functionality, etc.)	**Automation** (Ideal Type: IT systems are fully automated; i.e., only use pre-programmed verification of data entries and automated matching of data records)
Austria	0.75	1.00	1.00	0.75	0.75	0.49
England	0.49	0.25	0.25	0.25	0.25	0.25
Flanders	0.75	0.25	0.75	0.49	0.75	0.49
France	0.49	0.49	0.49	0.49	0.25	0.25
Germany	0.75	0.25	0.75	0.75	1.00	0.75
Greece	0.75	0.25	0.75	0.49	0.25	0.49
Hungary	0.49	0.25	0.25	0.49	0.49	0.49

Table 1: Fuzzy-membership scores for information quality and system quality components

4. CALIBRATING IMPACT

Using our existing knowledge on impacts, consequences or outcomes (e.g., Markus and Robey, 2004; Seddon, 1997) as a sensitising device, we extracted themes from our data that indicated positive impact or the lack thereof (e.g., error-free and timely data delivery, minimal or no duplication of work, satisfied stakeholders, etc.). Table 2 shows the codes that we drew out of the data through a dialogue between theoretical ideas and empirical evidence.

Concept	Code Group	Code Name
Impact		
	Intended consequences (2 codes; 10 open sub-codes such as, for example, questionable data validity, hidden costs, satisfied stakeholders, elimination of errors, etc.)	**Negative effects:** e.g., "One result of training is people that found a job after the end of the programme. On this, you do not need each time an evaluation study to have that, you have the monitoring system to tell you that six months after training that happened. This does exist as a monitoring system through the Centres for Vocational Training [i.e., a type of Beneficiary Organisation]. Yet the data validity in this is questionable as data is being entered by someone like the Centres for Vocational Training themselves, that is, it is in their interest to say that they did it! We then carried out an evaluation study and found out that some of the data was unreliable. Therefore, an evaluation study was required to show whether the data entered was valid or not. Unfortunately, this study translated in additional costs" (Head of ESF, Coordination Unit, Greece) **Positive effects:** e.g., "We are very satisfied with the filtering process embedded in the Templates. We find this filtering process very much adequate because accurate monitoring data is readily delivered to all concerned parties. This is very important to us because [this way] we can meet the ESF funding deadlines on time" (ESF Programme Manager, Ministry of Employment, State of Nortdrhein – Westphalia, Germany)
	Unintended consequences (2 codes; 23 open sub-codes such as, for example, duplication of work, delays in the delivery of data, seamless comparability of projects, monitoring as coaching, etc.)	**Negative effects:** e.g., "One single aligned [monitoring] system [between European and Flemish institutions] should benefit transparency and reduce costs. The main problem due to this misalignment is reflected in the definition of target groups. For example, the ESF Regulation establishes that the elderly are people above 45 whereas they are defined as those people above 50 by us. This means that we have to query our data set twice [over] to identify those people who are above 45 that are entitled to ESF funding. If the [European Commission] definitions were aligned with our data requirements it would be easier for us to meet them and less resource intensive" (Programme Manager, Partner Institution, ESF Agency, Flanders) **Positive effects**: e.g., "The fact that the monitoring arrangement is organised around the bilateral relation between Project Managers and Partner Institutions [i.e., Beneficiary Organisations] has introduced flexibility and confidence. The impression we have is that there is a good or deep knowledge of almost every project. We prefer the word 'coaching' as opposed to 'control' to refer to monitoring visits and audit controls" (Programme Manager, Partner Institution, ESF Agency, Flanders)

Table 2: Coding exemplars for key categories for impact

By combining these features, we conceived of our outcome of interest, i.e., positive impact, as the intended and/or unintended effects on organisational constituencies stemming from the enactment of the monitoring system. Though we used multiple sources to triangulate our data, we adopted the perspective of Managing/Paying Authorities and Monitoring Committees to assess these effects because of the mandatory nature of the monitoring system. Thus, we formulated judgments from the perspective of **these** stakeholders about what is beneficial from capturing and sharing validated monitoring data efficiently **and effectively**, and, conversely, what is not valuable from collecting, transmitting and validating monitoring data inefficiently **and ineffectively**. Subsequently, we derived 'ideal types' for positive cases and formulated summary statements that captured the evidence concerning every single case. Based on these statements, we rank-ordered each empirical case against its 'ideal type' and assigned fuzzy-membership scores accordingly (see Appendix 3).

For example, the following statement 'Little or no positive effect on relevant stakeholders' was used to characterise Greece. The deficiencies of the Integrated Data Warehousing System (e.g., summation issues in the context of physical data, systematic failures, etc.) are both resource and time consuming. As noted by the Head of the ESF Coordination Unit:

> "The Integrated Data Warehousing System allows us to easily follow what concerns the financial dimension. In what concerns the physical dimension, we need to receive [data] from the Managing Authorities. The Integrated Data Warehousing System is not a system that facilitates monitoring of the physical dimension… We use simple Microsoft Excel files as supporting tools to perform our function better. But these [tools] correspond to extra workload" [Coding category: negative unintended consequence].

These negative effects are magnified in the context of inconsistent business-level definitions between and among indicators. Again, the Head of the ESF Group for Monitoring and Evaluation said:

> "Many times, as the data is not ready [due to the lack of homogenisation], we spend time and man-hours in data collection and processing. If the indicator system were differently designed, we would save lots of time" [Coding category: negative unintended consequence].

Given the compounded negative effect stemming from technological deficiencies (e.g., additional workload, missing data due to systematic failures, etc.), delays in data aggregation, disaggregation and processing (i.e., late data delivery), lack of data accuracy and comparability (e.g., aggregation & disaggregation errors, undetected errors, traceability issues, etc.) and the opportunistic behaviours of Beneficiary Organisations (see 'negative intended consequence' entry in Table 2), we scored Greece 0.25 to emphasise that it is more out than in the set of cases with positive impact.

Conversely, the following statement 'Mostly positive effect on relevant stakeholders' was used to characterise Germany. The template system in Germany ensures a standardised accounting and reporting platform that compensates to a degree for the negative unintended effects stemming from the inconsistent set of result indicators (e.g., the inconsistent result indicators used across different States coupled with burgeoning participants' template data have undermined the aggregation and disaggregation of monitoring data). The template system, therefore, has facilitated a structured and rigorous approach for collecting, storing and transmitting monitoring data that ensures a coordinated monitoring strategy. As reported by an ESF Monitoring and Evaluation Officer:

> “The template system is a single accounting and reporting platform employing a uniform set of statistics, which the States, participating Ministries and the Federal Ministry of Labour & Social Affairs [i.e., the Federal Managing Authority] must adhere to. It has been introduced and implemented in such a way that it has resulted into a collaborative ground cultivated for all partners” [Coding category: positive intended consequence].

Even though the templates are time consuming to complete, Project Managers and Beneficiary Organisations have enjoyed intensive training sessions and helpdesk support to speed up the completion process in a standardised fashion. Again, as reported by an ESF Systems Manager:

> “We have provided information and training to Project Managers and Social Care Agencies [i.e., Beneficiary Organisations] on the use of the [template] system. Users were trained by us in conjunction with other organisations to align users’ understanding of concepts and terminology” [Coding category: positive intended consequence].

Furthermore, the intensive use of IT modules from the early stages of data entry enables Project Managers and Beneficiary Organisations to filter out individual sections of the templates, prompting early checks and timely delivery of monitoring data. As remarked by an ESF Programme Manager:

> “The [IT] system itself gives feedback. Firstly, feedback takes place when a project is being activated by the Project Manager. This gives the opportunity to the Project Manager to enter the data. The Project template itself covers more than 20 pages. During project data entry, individual sections may be filtered out, resulting to about 10 to 15 pages not used. Thereafter, however, the Project Manager needs to undertake the completion of Participants’ Templates which also through filtering out come to be about 10 pages each… Regarding the main uses of data, what is important is the Analysis module which is very helpful because, through the touch of button, we can filter out the state of individual projects in the course of [their] implementation at a certain point in time and check that the data is in good order to make sure that no errors occur during data input” [Coding category: positive intended consequence].

In addition, the seamless transmission of monitoring data ensures minimal data re-keying and, therefore, less room for errors. Given the overwhelming abundance of positive consequences whether intended (e.g., efficient coordination between state and federal monitoring activities, partial removal of discrepancies across result indicators, timely data delivery, etc.) or unintended (e.g., minimal duplication of verification procedures, few or no re-keying errors, etc.), we scored Germany 0.75 to emphasise that it is more in than out the set of cases with positive impact[2]. By repeating the same coding procedure for each case, we obtained the calibrations shown in Table 3.

Case/ Region	**Summary Statement**	**Assigned Fuzzy-Set Scores for Impact** (Ideal Type based on positive impact on relevant stakeholders only; e.g., error-free and timely delivery of required monitoring data with minimal use of resources, very satisfied stakeholders, etc.)
Austria	Mostly positive effect on relevant stakeholders thanks to a concerted monitoring approach revolving around a consistent set of indicators (e.g., seamless comparability of training projects in terms of achieved targets, satisfied stakeholders, etc.). However, manual validations have triggered occasional delays in terms of data delivery, some data processing errors and duplication of manual controls (i.e., unnecessary costs).	0.75
England	Little or no positive effect on relevant stakeholders because of conflicting perceptions of information requirements (e.g., widespread lack of satisfaction across stakeholder organisations worsened by out-of-date indicators), manual data re-keying (e.g., numerous re-keying errors, duplication of costs, etc.) and manual validations of inconsistent data (e.g., undetected errors, data delivery delays, etc.).	0.25
Flanders	Instances of both negative effects (e.g., duplication of work stemming from submitting reports in accordance with both ESF and Flemish definitions) and positive effects (e.g., data convergence documents facilitating seamless validation of CFS data and its timely delivery). Quality control reports may trigger undetected errors because they are based on manual validations. Yet they are an occasion for monitoring visits and audit controls that nurture good bilateral relations between Project Managers and Partner Institutions, thus generating deep knowledge of ESF projects.	0.49
France	Little or no positive effect on relevant stakeholders stemming	0.25

[2] Germany has automated the monitoring process to speed up the validation and transmission of monitoring data because entering a large number of data items in the templates was too time consuming. Hence, the efficiency savings in terms of lower validation costs and fewer re-keying errors were largely unintended.

	from monitoring requirements perceived to be too bureaucratic and not geared towards a common purpose (e.g., lack of satisfaction across National and Regional Managing Authorities). Manual validations of inconsistent data generate data checking errors. There are also data losses stemming from frequent breakdowns, undetected errors caused by softer certification issues, as well as unnecessary delays in the delivery of monitoring data and re-keying errors (e.g., high-error rates due to re-keying of financial data from regional databases, delays due to manual data re-keying, etc.).	
Germany	Mostly positive effect on relevant stakeholders thanks to a standardised approach for the collection, storage and transmission of monitoring data leading to more efficient coordination between Federal and State monitoring activities and to seamless verification and transmission of monitoring data (e.g., timely data delivery, few re-keying errors, minimal duplication of costs, etc.). However, overwhelming data requirements and inconsistent result indicators have led to minor errors in the aggregation & disaggregation of monitoring data at the Federal level.	0.75
Greece	Little or no positive effect on relevant stakeholders stemming from technical limitations of the Integrated Data Warehousing System causing additional workload (i.e., unnecessary costs) and systematic failures (i.e., data losses leading to missing data). Manual validations of inconsistent physical data trigger undetected errors and data processing delays. There are also additional monitoring costs because of Beneficiary Organisations' opportunistic behaviours.	0.25
Hungary	Little or no positive effect on relevant stakeholders stemming from inconsistent and out-of-date indicators (e.g., poor traceability of physical output & result data, low satisfaction levels across sector-specific Managing Authorities, etc.), lack of interoperability across interfaces (e.g., time-consuming data inputs, manual extraction of data, etc.), slow IT systems unable to save tables properly (e.g., data losses) and manual validations of textual data (e.g., undetected	0.25

	errors when it comes to textual field checks).	

Table 3: Fuzzy-membership scores for Impact

5. ANALYSIS

This Section presents the fsQCA analysis of the cases described above, conceiving of them as configurations (or combinations) of explanatory conditions (Fiss, 2011). In coding and calibrating the data, we combine sets of causal components that go together conceptually. These are then re-conceptualised at a higher level of abstraction in terms of information and system quality in accordance with the original DeLone and McLean model. We now assess the degree of membership in each higher-order concept using the rule of the minimum to check whether individual cases conform to 'ideal types'. This rule assumes that a case with weak membership in one of its higher-order concepts' components could have, at best, only weak membership in any combinations of parts that include this specific component (Ragin, 2008). Table 4 shows the resulting fuzzy-membership scores for information quality and system quality.

Country Case	**Fuzzy-membership scores for Information Quality**	**Fuzzy-membership scores for System Quality**
Austria	0.75	0.49
England	0.25	0.25
Flanders	0.25	0.49
France	0.49	0.25
Germany	0.25	0.75
Greece	0.25	0.25
Hungary	0.25	0.49

Table 4: Aggregated fuzzy-membership scores for Information Quality and System Quality

We then create a data matrix, Table 5, to outline the distribution of cases across the four logically-possible combinations of aggregated causal conditions.

Country/Region	IQ*SQ	iq*SQ	IQ*sq	iq*sq
Austria	0.49	0.25	**0.51**	0.25
England	0.25	0.25	0.25	**0.75**
Flanders	0.25	0.49	0.25	**0.51**
France	0.25	0.25	0.49	**0.51**
Germany	0.25	**0.75**	0.25	0.25
Greece	0.25	0.25	0.25	**0.75**
Hungary	0.25	0.49	0.25	**0.51**
Number of cases with score > 0.50 in bold	0	**1**	**1**	**5**

Legend:
iq= Non-Information Quality;
IQ= Information Quality;
sq= Non-System Quality;
SQ= System Quality;
Logical AND designated by the asterisk (*) to stipulate the intersection of sets

Table 5: Distribution of cases across combinations of aggregated causal conditions

Next, we create a truth table that lists all logically-possible combinations of aggregated causal conditions whether they are populated with cases or not (see Table 6). Setting the minimum number-of-cases threshold at 1 and establishing the lowest level of acceptable consistency at 0.85, Table 6 shows as fuzzy subsets of the outcome those combinations of aggregated causal conditions whose consistency score is at or above the cut-off value of 0.85.

IQ	SQ	Number of cases with score > 0.5	Outcome code based on consistency score	Raw Consistency	PRI Consistency	SYM Consistency
0	0	5	0	0.56	0.00	0.00
0	1	1	1	0.91	0.67	0.67
1	0	1	1	0.89	0.52	0.52
1	1	0	Remainder			

Legend:
Raw Consistency= degree to which the combination of causal conditions is a subset of membership in the outcome (the default way of assessing consistency)
PRI Consistency= Proportional Reduction Interpretation Consistency (an additional parameter to interpret set-theoretic relations)
SYM Consistency= Symmetry Consistency (a tweaked version of PRI Consistency)
Remainders designate configurations with no empirical cases with membership scores above 0.50
0= absence; 1= presence (only applied to IQ and SQ values)

Table 6: Truth table

By inspection, Table 6 shows that two configurations are sufficient for positive impact, namely the combination of non-Information Quality AND System Quality and the combination of Information Quality and non-System Quality (see rows 2 & 3). At this juncture, there are two possibilities to pursue. The first and more conservative option is to avoid 'remainders' altogether and stick with this solution. However, in-depth knowledge of the empirical data shows that Germany is far from being a good instance of non-Information Quality and System Quality because, although it is closer to the iq*SQ corner of the vector space, its standardised approach for the collection, storage and transmission of monitoring data compensates to a degree for the fragmentation of result indicators. In other words, it is much easier to aggregate fragmented physical result data if it is stored and transmitted to the Federal Managing Authority in a standardised fashion through the 'template system'. Likewise, Austria is not an exemplary instance of Information Quality and non-System Quality because the use of a consistent set of indicators compensates to a degree for the lack of automated validations thanks to the 'four-eye' principle (as pointed out by the fuzzy score in Table 5). Thus, based on our in-depth case knowledge, we choose to move beyond the constraints of this complex solution. The alternative is to incorporate the 'remainder' row, that is, IQ AND SQ (IQ * SQ), in order to yield a more parsimonious solution. Table 7 shows the incorporation of the remainder row.

iq*SQ + IQ*SQ + IQ*sq ⟶ Impact Where: iq= Non-Information Quality; IQ= Information Quality; sq= Non-System Quality; SQ= System Quality **Legend**: Logical AND designated by the asterisk (*); Logical OR designated by the plus sign (+) Sufficiency designated by the arrow running from the sufficient conditions towards the outcome (⟶)

Table 7: Incorporation of remainder row

By using Boolean algebra (Cf. Ragin, 2008: 156), Table 8 arrives at the parsimonious solution.

SQ (iq + IQ) + IQ (SQ + sq) ⟶ Impact SQ + IQ ⟶ Impact

Table 8: Parsimonious solution

From here we go on to uncover more granular and temporally-meaningful causal configurations (or solutions) within these scope conditions, acknowledging that scope conditions are claims about the domain in which causal effects are stable and that one cannot arbitrarily raise the number of conditions regardless of the number of cases (Marx, 2010). Using process tracing, we analyse the evidence on processes, sequences and conjunctures of events within the cases as the dialogue between theoretical and substantive knowledge is pursued. Thus, by returning to the cases, we can unravel several temporal pathways leading to the outcome of interest (and its absence). Figure 5 depicts such pathways.

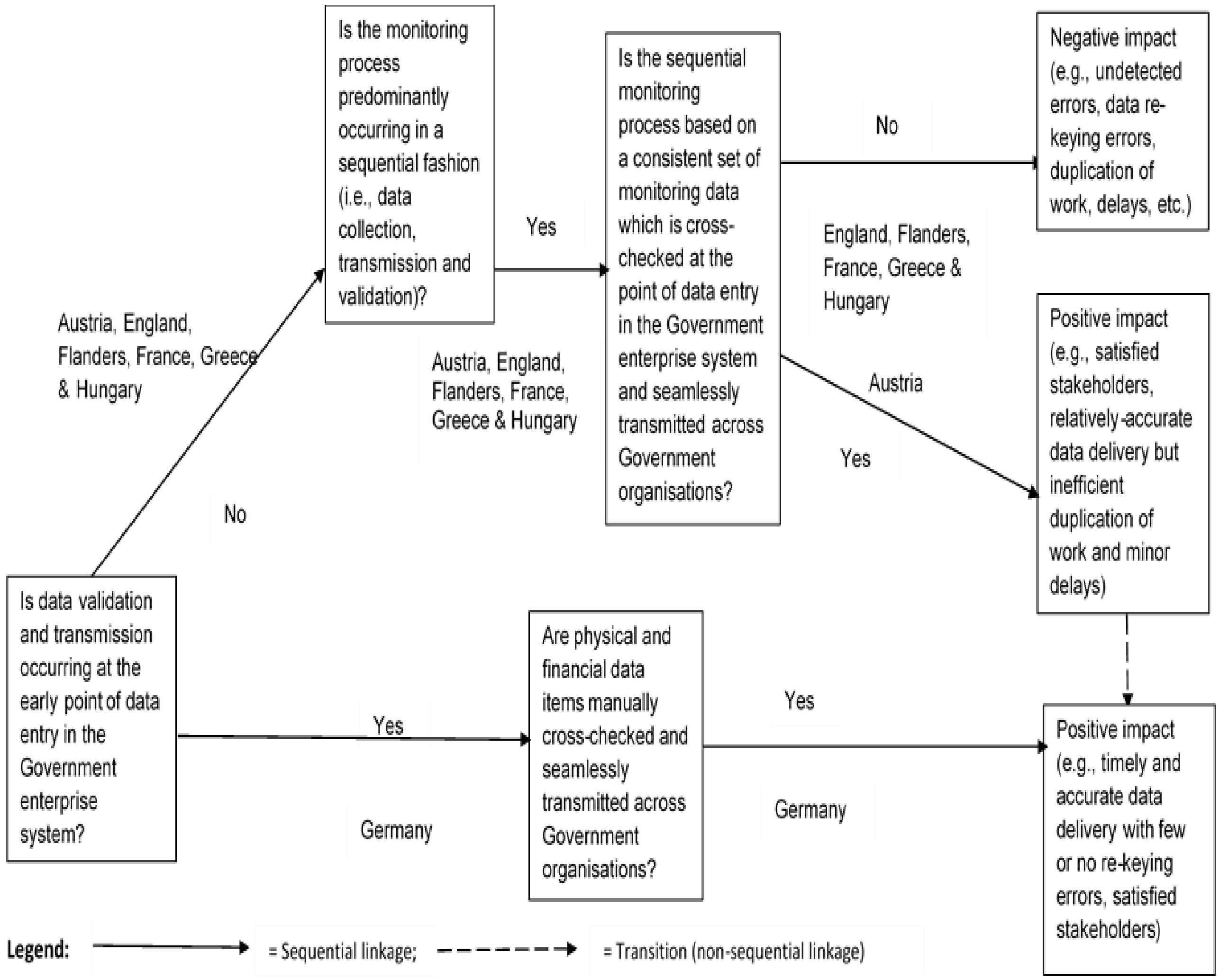

Figure 5: Temporal pathways to positive and negative impact in the 2000-2006 programming period. Adapted from George and Bennett (2005).

The pathways shown in Figure 5 correspond to two distinct archetypes of successful monitoring systems. One pathway is characterised by a seamless validation and transmission of monitoring data occurring at the early point of data entry (i.e., Germany). Another pathway reflects a ‘hybrid’ configuration which is on a journey from manual to automated validation (i.e., Austria). Accordingly, we interpret Germany to be a good instance of a ‘well-oiled electronic machine’ and Austria to be a good instance of a ‘hybrid machine’ where manual validation revolves around the ‘four-eye’ principle but based on a consistent set of indicators. By pinpointing the causal conditions mobilised in these two pathways and factoring out the commonalities between these two types of ‘electronic bureaucracies’, we obtain two distinct configurations (Cf. Berg-Schlosser et al., 2009: 16). Table 9 shows the causal configurations underpinning each archetype.

	SOLUTIONS[a]	
Configurations	**S1 (Well-oiled electronic machine/Germany)**	**S2 (Hybrid machine/Austria)**
Comprehensiveness	●	●
Consistency		•
Currency	●	●
Compatibility	●	●
Reliability	●	●
Automation	•	
Consistency (sufficiency)	**1.00**	**1.00**
Raw coverage	**0.91**	**0.75**
Unique coverage	**0.25**	**0.09**
Overall solution consistency	**1.00**	
Overall solution coverage	**1.00 (out of which 0.66 is overlapping coverage)**	

Table 9: Causal configurations for achieving positive impact. Adapted from Fiss (2011).
[a] Black circles indicate the presence of a condition. Large circles indicate core (necessary) conditions; small ones, peripheral (contingent) conditions. Blank spaces indicate 'don't care', that is, situations where causal conditions may be either present or absent.

Next, we identify the causally-relevant commonalities between Germany and Austria as the causal core and designate the remaining causes as being part of the periphery. This reveals two configurations for positive impact. To achieve an efficient and effective monitoring process, it is necessary that a monitoring system collects a comprehensive and up-to-date range of financial and physical indicators and relies on dependable technologies with well laid out communication protocols. However, this is not enough. Raw monitoring data must be validated to verify that it has been entered correctly in the system whether at the individual or aggregate level. Two procedures seem possible for this purpose: either validation in an automated fashion or manual validation. While the former may guarantee a speedy and seamless processing of monitoring data at the early stages of the monitoring process, the latter puts a strain on speed and accuracy because human beings are slower and less accurate than technologies in such information processing tasks. As remarked by one Programme Manager in Austria:

> "Prior to the introduction of the SAP module going live as part of the Data Warehousing System in July 2005, some errors were occurring that were found

to be associated with manual procedures concerning data processing and entry. In the new system platform, these errors have been eliminated, whilst there is still some accounting workload on projects and actions that were completed before July 2005".

Hence, Member States using manual validation procedures must deploy a consistent set of indicators if they are to keep up with the speed and accuracy that can be achieved by technological automation. The Austrian case is telling in this respect. Data transmitted to Beneficiary Organisations is manually cross-checked by at least two colleagues per office through the 'four-eye' principle and, once entered in the Data Warehousing System, it is sent to the Federal Managing Authority where it undergoes further checks. The overall efficiency of these manual validation procedures depends on consistent definitions of monitoring data and indicators. The more consistent the set of indicators, the faster and more accurate the manual checking and cross-checking of data entries can be.

Conversely, the lack of positive impact (i.e., the failure of the monitoring system) presupposes no compatible communication protocols coupled with inconsistent indicators and **non-automated (i.e., manual)** validations. In other words, the manual verification of inconsistent data in conjunction with the re-keying of data entries is a guaranteed **configuration** for failure of the monitoring system. Table 10 shows the causal configuration for achieving negative impact.

Configurations	SOLUTION[a] S3 (Stuttering electronic machine; e.g., Flanders)
Comprehensiveness	
Consistency	⊗
Currency	
Compatibility	⊗
Reliability	
Automation	⊗
Consistency (sufficiency)	1.00
Raw coverage	0.76
Unique coverage	0.76
Overall solution consistency	1.00
Overall solution coverage	0.76

Table 10: Causal configuration for achieving negative impact. Adapted from Fiss (2011).
[a] Circles with "x" indicate the absence of a condition. Large circles indicate core (necessary) conditions. Blank spaces indicate "don't care", that is, situations where causal conditions may be either present or absent.

Again, this configuration constitutes a different archetype of monitoring system that we label a 'stuttering electronic machine' because of the frequent re-keying of monitoring data, unwarranted duplication of work, undetected errors and delays.

6. DISCUSSION AND CONCLUSION

The DeLone and McLean model has broken new ground in the study of IS success **helping uncover** common threads across a variety of studies. Yet, it is based on very simple assumptions about causality and temporality. **Specifically, it is not clear why the (causal) relationships between independent variables (or causal conditions) and dependent variable (or outcome of interest) are simply linear rather than being contingent on moderating variables (Burton-Jones et al., 2015). Likewise, it is not clear why the model assumes a straightforward chain of events where "information flows through a series of stages from its production through its use or consumption to its influence on individual and/or organizational performance" (DeLone and McLean, 1992: 61).** Spurred by calls for further research to re-invigorate debate on the quest for the dependent variable (Burton-Jones et al., 2015; Dwivedi et al., 2015), we asked at the outset of this paper: 1) What aspects of the ESF monitoring system are relevant for a positive effect (i.e., outcome, consequence or impact) to be seen? 2) **In what ways do** such aspects produce these effects?

We are now in a position to answer these questions. Specifically, we argue that there are two distinct aspects that **produce** positive effects, namely temporal and configurational aspects. From **the** temporal perspective, there are **sequences** that lead to positive downstream effects, each stage of which is necessary but not sufficient for the final outcome. In particular, we have identified two separate sequences leading to positive impact depending on whether data collection, validation and transmission are mostly co-occurring events or are **straightforward chains of events where data collection precedes the electronic transmission of data to Beneficiary Organisations which, in turn, precedes manual validation via the 'four-eye' principle (see Figure 5)**. From a configurational perspective, each temporal pathway triggers different combinations of causal conditions which share several 'core' (or necessary) conditions. Thus, we see this paper as re-appraising the causal and temporal influences underpinning IS success in the ESF context, thereby contributing to theory, methodology and practice.

Theoretically, we demonstrate that the relationships between causal conditions and the outcome of interest need not be treated as crudely linear (e.g., "if more X then more Y". Cf. Markus and Robey, 1988: 590). On the contrary, causal conditions work together rather than separately because they are **parts of more complex systems** that produce holistic effects not inherent in their individual parts (Mahoney, 2001). Causal conditions may complement each other in some contexts and counteract one another in other contexts depending on the underpinning interaction mechanism. For example, compatibility counteracted the lack of consistency in Germany thanks to the exchange of structured data across the interface (i.e., the 'template system'). Likewise, consistency compensated for the lack of automation in Austria thanks to a concerted monitoring approach relying on a set of consistent indicators and the use of the 'four-eye' principle. While both these countries experienced a positive downstream impact, other countries were not so fortunate. For example, in England, Flanders, France, Greece and Hungary the lack of consistency of monitoring data coupled with the lack of automated validation and the absence of shared communication protocols caused data delivery delays, numerous errors (e.g., data re-keying errors, undetected errors, etc.), as well as additional overheads in terms of duplication of work.

Methodologically, the use of **Cornford's et al. (1994) structure-process-outcome framework** has allowed us to re-consider the process dimension of the DeLone and McLean model. Although this model is based on a sequential understanding of causality, our analysis has shown two equifinal pathways to the outcome of interest. Not only do these pathways 'trigger' different configurations of causal conditions. They may also embody co-occurring events rather than simple **chains of events** (or sequences). For example, in Germany data validation and transmission occur to a large extent simultaneously at the early point of data entry. Recently, scholars have called for new methods for tracing the processes and sequential links that underpin IS phenomena of interest (Avgerou, 2013). **Our typological approach offers some progress in this direction. Specifically, the integration between fsQCA and process tracing allows for structured iterations between theory and cases, thus linking theoretical and empirical strands more closely together (George and Bennett, 2005). Indeed, the dialogue between theory and empirical data plays a crucial role in this respect because it fosters the development of concepts that are not defined at too high a level of abstraction.** Thus, in this paper, we have cycled back and forth between theory and empirical data to **develop** concepts (and their associated **components**) that are not detached from the context under investigation. Accordingly, we have developed a typological theory that lies midway between the temporal pathways embedded within individual cases and broader theoretical patterns emerging across cases.

Typological theories consist of configurations of core and peripheral parts, with the core elements being essential (or necessary) and the peripheral elements being less important and perhaps even "expendable or exchangeable" (Fiss, 2011: 394). **They enable scholars to regard** IS success as a concept that is neither additive nor multiplicative (Polites et al., 2012) but rather substitutable, where one type of successful monitoring system can substitute for another offering a revised temporality. So, in our cases, while 'hybrid machines' may substitute for 'well-oiled electronic bureaucracies', it turns out that the latter are a more powerful type because they foreshadow the move towards monitoring systems acting as "early-warning systems" (Tödtling-Schönhofer et al., 2011: 13/26)[3].

There are also implications for practice in this research, in particular in relation to issues of asymmetry (Fiss, 2011; Liu et al., 2015). For managers and policy makers alike the insight that configurations leading to monitoring systems success are not simply the reverse of the configurations for failure may have great value. While successful monitoring systems presuppose dependable technologies with well-defined communication protocols able to capture and transmit a comprehensive and up-to-date range of indicators, unsuccessful monitoring systems are characterised in a different way. They show a lack of socio-technical standards and non-automated (i.e., manual) validations. Thus, the removal of a required condition for failure may be desirable and necessary, but is hardly sufficient for success.

In simple words, managing for non-failure is not going to achieve success. But, in any given situation, managing for success may go beyond the resources available, so careful choices need to be made. For a manager or policy maker, a typological theory such as developed here (see Table 9) can offer a way through this conundrum, helping to make a better judgement as to which path to take – put very simply whether to pursue manual or automated validations and by what means. In this way, the analysis here might tell a manager or policy maker that manual validations relying on consistent indicators can substitute for the lack of automated checks when, for example, using the 'four-eye' principle, but may have resource implications. Conversely, automated validations too may need manual cross-checks to ensure that data entered in the templates matches its original sources.

6.1 Limitations and future research directions

This work has limitations and offers a number of avenues for future research. Here we address what we see as the four most significant issues chosen both because they link with our specific findings regarding monitoring systems, but also because they point to important lessons that may inform IS scholars following, methodologically, in our footsteps.

First, and directly linked to the practical implications, we have regarded our cardinal concept, success, and its opposite, failure, as symmetric concepts (Goertz and Mahoney, 2012). Thus, we have assumed that the negated concept, i.e., the lack of success, is the same

---

[3] The raw coverage (a gauge of empirical importance) corroborates this argument as it is larger for Germany than Austria (see Table 9).

thing as the opposite concept, i.e., failure, and that, therefore, the failure of the monitoring system corresponds to its negative impact or, symmetrically, that its success literally is its positive impact. However, as Schneider and Wagemann (2012) suggest, it would have been preferable to re-think the selection of causal conditions and construct a new truth table for unsuccessful monitoring systems rather than assuming that the same conditions affect both monitoring systems success and failure. More generally, for any other IS scholar taking this route, we suggest running a separate analysis for the negation of the outcome to better grasp the causal logic for positive cases, but then to address the analysis of negative cases in a follow-up study by considering different theories (and conditions) that are specifically tailored to the non-occurrence of the outcome. In our case, this would imply moving from the DeLone and McLean model of IS success to some failure model (Davis et al., 1992; Dwivedi et al., 2013; Lyytinen and Robey, 1999; **Tan et al., 2016**).

Second, although Basurto and Speer's (2012) calibration procedure can be applied using either a deductive or a "more open, inductive research strategy" (Ibid: 160), language and cultural barriers in our international research team may have hindered the open-coding process underpinning the inductive application of their approach. Again, IS scholars wishing to adopt our methods may instead use fsQCA deductively. For example, they could deductively pinpoint several dimensions and ask country experts to group them into various constructs' categories (or sets) which could then be calibrated by means of explicit rules (Verkuilen, 2005). These calibrated constructs could then be used to create a logical structure of possibilities before studying individual cases (George and Bennett, 2005).

Third, our findings should be taken with care because we acknowledge that we have relied primarily on country experts' specialised knowledge which can bias the calibrations of data and, indirectly, the precision of our findings (McGrath, 1981). Researchers wishing to follow in our footsteps could seek to reduce such biases by triangulating their data sources more systematically than we have done.

Fourth, and last, the data analysed in this paper refers to the 2000-2006 programming period rather than the more recent 2007-2013 or 2014-2020 periods. The reasons for this is simply that data was initially collected in 2006-2008 based on a desire to explore and learn from the levels of monitoring systems success achieved in this period. Still, this time lag does raise the issue of predictive validity. We have not gathered or analysed more recent primary data, but our findings may be expected to hold in the more recent programming periods since the ESF is still predicated upon the same schema introduced by Regulation 1260/99 in terms of physical and financial indicators (Cf. European Commission, 2006: 5-7; 31-32; European Commission, 2015: 9 – 10; 28-30; Regulation 1303/2013: Article 27).

These limitations notwithstanding, we contend that this paper has re-appraised the causal and temporal influences underpinning IS success in the ESF context. **Using a typological approach**, we have unravelled the causal mechanisms and temporal pathways underpinning monitoring systems success. On this basis, we have developed a typological theory that unpacks several types and mixed types of monitoring systems **in the ESF context**. Although they are spatially and temporally-bounded, we see typological theories as offering a valuable way of bridging the ever-growing gap between theory and practice in the IS field.

REFERENCES

Avgerou, C. (2013) Social mechanisms for causal explanation in social theory based IS research. Journal of the Association for Information Systems, 14, 399-419.

Ballantine, J., Bonner, M., Levy, M., Martin, A., Munro, I. & Powell, P.L. (1996) The 3-D model of information systems success: the search for the dependent variable continues. Information Resources Management Journal, 9, 5-15.

Basurto, X. & Speer, J. (2012) Structuring the calibration of qualitative data as sets for qualitative comparative analysis. Field Methods, 24, 155-174.

Berg-Schlosser, D., De Meur, G, Rihoux, B. & Ragin, C. (2009) Qualitative Comparative Analysis (QCA) as an Approach. In: Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques, Rihoux, B. and Ragin, C. (Eds.), pp. 1-18, Sage Publications, London.

Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology. Qualitative Research in Psychology, 3, 77-101.

Burton-Jones, A., McLean, E. R. & Monod, E. (2015) Theoretical perspectives in IS research: from variance and process to conceptual latitude and conceptual fit. European Journal of Information Systems, 24, 664-679.

Caldeira, M. M. & Ward, J. M. (2002) Understanding the successful adoption and use of IS/IT in SMEs: an explanation from Portuguese manufacturing industries. Information Systems Journal, 12, 121-152.

Cornford, T., Doukidis, G. & Forster, D. (1994) Experience with Structure, Process and Outcome Framework for Evaluating an Information System. Omega International Journal of Management Science, 22: 491-504.

Cress, D. M. & Snow, D. A. (2000) The outcomes of homeless mobilization: The influence of organization, disruption, political mediation, and framing. American Journal of Sociology, 105, 1063-1104.

Crilly, D. (2011) Predicting stakeholder orientation in the multinational enterprise: A mid-range theory. Journal of International Business Studies, 42, 694-717.

Crilly, D., Zollo, M. & Hansen, M. T. (2012) Faking it or muddling through? Understanding decoupling in response to stakeholder pressures. Academy of Management Journal, 55, 1429-1448.

Davis, G. B., Allen, L. S., Nickles, K. R., Chatterjee, S., Hartung, R.& Wu, Y. (1992) Diagnosis of an information systems failure: a framework and interpretive process. Information and Management, 23, 293-318.

DeLone, W. H. & McLean, E. R. (1992) Information Systems Success: The Quest for the Dependent Variable. Information Systems Research, 3, 60-95.

DeLone, W. H. & McLean, E. R. (2003) The DeLone and McLean model of information systems success: a 10 year update. Journal of Management Information Systems, 19, 9-30.
DeLone, W. H. & McLean, E.R. (2004) Measuring e-commerce success: Applying the DeLone & McLean information systems success model. International Journal of Electronic Commerce, 9, 31-47.
DeLone, W. H. & McLean, E. R. (2016) Information Systems Success Measurement. Foundations and Trends® in Information Systems, 2, 1-116.
De Meur, G., Rihoux, B. & Yamasaki, S. (2009) Addressing the critiques of QCA. Qualitative Comparative Analysis (QCA) as an Approach. In: Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques, Rihoux, B. and Ragin, C. (Eds.), pp. 1-18, Sage Publications, London.
Dwivedi, Y. K., Wastell, D., Laumer, S., Henriksen, H. Z., Myers, M. D., Bunker, D. , Elbanna, A., Ravishankar, M. N. & Srivastava, S. C. (2015) Research on information systems failures and successes: Status update and future directions. Information Systems Frontiers, 17, 143-157.
Dwivedi, Y. K., Ravichandran, K., Williams, M. D., Miller, S., Lal, B., Antony, G. V. & Kartik, M. (2013) IS/IT Project Failures: A Review of the Extant Literature for Deriving a Taxonomy of Failure Factors. In: Grand Successes and Failures in IT: Public and Private Sectors, Dwivedi, Y. K., Henriksen, H. Z., Wastell, D. & De', R. (Eds.), pp. 73-88. Springer, London.
El Sawy, O., Malhotra, A, Park, Y. & Pavlou, P. A. (2010) Seeking the configurations of digital ecodynamics: it takes three to tango. Information Systems Research, 21, 835-848.
European Commission: European Regulation 1260/99. [WWW document] Available at: http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:1999:161:0001:0042:EN:PDF (Access date: 26/07/2016).
European Commission: Directorate-General, Regional Policy (2006) The New Programming Period 2007-2013. Indicative Guidelines on Evaluation Methods: Monitoring and Evaluation Indicators. [WWW document] Available at: http://ec.europa.eu/regional_policy/sources/docoffic/2007/working/wd2indic_082006_en.pdf (Access date: 26/07/2016).
European Commission: European Regulation 1303/2013. [WWW document] Available at: http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R1303&from=EN (Access date: 26/07/2016).
European Commission: Employment, Social Affairs & Inclusion (2015) Programming period 2014-2020. Monitoring and evaluation of European Cohesion Policy, European Social Fund. Guidance document. [WWW document] Available at: http://ec.europa.eu/sfc/en/system/files/ged/Annex%20D%20%E2%80%93%20Practical%20

guidance%20on%20data%20collection%20and%20validation.pdf (Access date: 26/07/2016).
Fiss, P. C. (2011) Building better causal theories: a fuzzy set approach to typologies in organization research. Academy of Management Journal, 54, 393- 420.
Gable, G. G., Sedera, D., & Chan, T. (2008) Re-conceptualizing information system success: The IS-impact measurement model. Journal of the Association for Information Systems, 9, 377-408.
George, A., & A. Bennett (2005) Case studies and theory development in the social sciences. MIT Press, Cambridge, Massachusetts.
Goertz, G. & Mahoney, J. (2012) A tale of two cultures: qualitative and quantitative research in the social sciences. Princeton University Press, Princeton, NJ.
Henfridsson, O., & Bygstad, B. (2013) The generative mechanisms of digital infrastructure evolution. MIS Quarterly, 37, 907-931.
Klecun, E., Lichtner, V., Cornford, T. & Petrakaki, D. (2014) Evaluation as a Multi-Ontological Endeavour: A Case from the English National Program for IT in Healthcare. Journal of the Association for Information Systems, 15, 147-176.
Lamb, R. & Kling R. (2003) Reconceptualising users as social actors in information systems research. MIS Quarterly, 27, 197-235.
Lee, A. (2010) Retrospect and prospect: information systems research in the last and next 25 years. Journal of Information Technology, 25, 336-348.
Liu, Y., Mezei, J., Kostakos, V. & Li, H. (2017) Applying configurational analysis to IS behavioural research: a methodological alternative for modelling combinatorial complexities. Information Systems Journal, 27, 59-89.
Lyytinen, K. & Robey, D. (1999) Learning failure in information systems development. Information Systems Journal, 9, 85-101.
Mahoney, J. (2001) Review essay: Beyond correlational analysis: recent innovations in theory and method. Sociological Forum, 16, 575-593.
Markus, L. M.& Robey, D. (1988) Information technology and organizational change: causal structure in theory and research. Management Science, 34, 583-598.
Markus, L. M. & Robey, D. (2004) Why Stuff Happens: Explaining the Unintended Consequences of Using Information Technology. In: The Past and Future of Information Systems, Vendelo, M. T. and Andersen, K. V. (Eds), pp. 61-93, Elsevier Butterworth-Heinemann, Amsterdam, Netherlands.
Marx, A. (2010) Crisp set qualitative comparative analysis (csQCA) and model specification: benchmarks for future csQCA applications. International Journal of Multiple Research Approaches, 4, 138-158.

Mason, R. O. (1978) Measuring information output: A communication systems approach. Information & Management, 1, 219-234.

McGrath, J. E. (1981) Dilemmatics: The Study of Research Choices and Dilemmas. American Behavioral Scientist, 25, 179-210.

Mingers, J. (2014) Systems thinking, critical realism and philosophy: A confluence of ideas. Routledge, London.

Mumford, E. (2006) The story of socio-technical design: reflections on its successes, failures and potential. Information Systems Journal, 16, 317–342.

Park, Y. & El Sawy, O. A. (2013) The value of configurational approaches for studying digital business strategy. In: Configurational Theory and Methods in Organizational Research, Cambré , B., Fiss, P. C. and Marx, A. (Eds.), pp. 205-224, Emerald Group Publishing Ltd, Bingley.

Petter, S., DeLone, W. & McLean, E. R. (2008) Measuring information systems success: models, dimensions, measures, and interrelationships. European Journal of Information Systems, 17, 236-263.

Petter, S., DeLone, W. & McLean, E. R. (2012) The past, present and future of IS "success". Journal of the Association for Information Systems, 13, 341-362.

Petter, S., DeLone, W. & McLean, E. R. (2013) Information Systems Success: the quest for the independent variables. Journal of Management Information Systems, 29, 7-62.

Polites, G., L., Roberts, N.& Tatcher, J. (2012) Conceptualising models using multidimensional constructs: a review and guidelines for their use. European Journal of Information Systems, 21, 22-48.

Ragin, C. C. (1987) The comparative method: moving beyond qualitative and quantitative strategies. University of California Press, Berkeley.

Ragin, C. C. (2000) Fuzzy set social science. The University of Chicago Press, Chicago.

Ragin, C. C. (2008) Redesigning Social Enquiry. The University of Chicago Press, Chicago.

Rihoux, B., & Lobe, B. (2009) The Case for Qualitative Comparative Analysis (QCA): Adding Leverage for Thick Cross-Case Comparison. In: The Sage Handbook of Case-Based Methods, Byrne, D. & Ragin, C. C. (Eds.), pp. 222-242. Sage Publications Ltd., London.

Rivard, S. & Lapointe, L. (2012) Information Technology Implementers' Responses to User Resistance: Nature and Effects. MIS Quarterly, 36, 897-920.

Rosemann, M. & Vessey, I. (2008) Toward improving the relevance of information systems research to practice: The role of applicability checks. MIS Quarterly, 32, 1–22.

Schneider, C. Q., & Wagemann C. (2012) Set-Theoretic methods for the social sciences: a guide to qualitative comparative analysis. Cambridge University Press, Cambridge.

Scott, M., DeLone, W. & Golden, W. (2016) Measuring eGovernment success: a public value approach. European Journal of Information Systems, 25, 187-208.

Seddon, P. B. (1997) A respecification and extension of the DeLone and McLean model of IS success. Information Systems Research, 8, 240-253.
Shannon, C. E., & Weaver, W. (1949) The Mathematical Theory of Communication. University of Illinois Press, Urbana, IL.
Smithson, S., & Hirschheim, R. (1998) Analysing information systems evaluation: another look at an old problem. European Journal of Information Systems, 7, 158-174.
Tan, C. W., Benbasat, I., & Cenfetelli, R. T. (2016) An Exploratory Study of the Formation and Impact of Electronic Service Failures. MIS Quarterly, 40, 1-29.
Tödtling-Schönhofer, H., Naylon, I., Génin, H., Basile, I. & Savioli, B. (2011) Synthesis Report Metis GmbH in cooperation with Euréval. [WWW document] Available at: http://ec.europa.eu/social/BlobServlet?docId=7342&langId=en (Access date: 26/07/2016).
Urbach, N., Smolnik, S. & Riempp, G. (2009) The state of research on information systems success - a review of existing multi-dimensional approaches. Business & Information Systems Engineering, 4, 315-325.
Verkuilen, J. (2005) Assigning membership in a fuzzy set analysis. Sociological Methods & Research, 33, 462-496.

SUPPORTING INFORMATION FOR REVIEW AND ONLINE PUBLICATION ONLY

The list of key sources and activities is reported in Appendix 1. The core features of the seven cases under investigation are described in Appendix 2. Fuzzy-set calibration thresholds are reported in Appendix 3. Fuzzy-membership scores for information and system quality components with their associated summary statements are presented in Appendix 4. Coding exemplars for key features of information and system quality components are outlined in Appendix 5. The interview guide, primary and secondary data, as well as the summary of calibrations are available upon request.