



Universiteit
Leiden
The Netherlands

Understanding evaluation use from an organisational perspective: a review of the literature and a research agenda

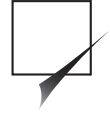
Kupiec, T.; Celinska-Janowicz, D.; Pattyn, V.E.

Citation

Kupiec, T., Celinska-Janowicz, D., & Pattyn, V. E. (2023). Understanding evaluation use from an organisational perspective: a review of the literature and a research agenda. *Evaluation-Us*, 29(3), 338-355.
doi:10.1177/13563890231185164

Version: Publisher's Version
License: [Creative Commons CC BY 4.0 license](#)
Downloaded from: <https://hdl.handle.net/1887/3716487>

Note: To cite this publication please use the final published version (if applicable).



Article

Understanding evaluation use from an organisational perspective: A review of the literature and a research agenda

Evaluation

2023, Vol. 29(3) 338–355

© The Author(s) 2023



Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/13563890231185164

journals.sagepub.com/home/evi



Tomasz Kupiec 
University of Warsaw, Poland

Dorota Celińska-Janowicz 
University of Warsaw, Poland

Valérie Pattyn 
Leiden University, The Netherlands

Abstract

Why do some types of evaluation use prevail in certain contexts and not in others? The aim of this article is to advance knowledge about organisational factors of evaluation use, that is, determinants of evaluation use grounded in organisational theories. We critically review existing frameworks of organisational factors of evaluation use, highlighting key differences between them and pointing out discrepancies with empirical insights. We discuss the merits of two potential areas for future research that can help concretise theoretical stances: considering organisational legitimacy as a potential direct determinant of evaluation use and incorporating a dynamic perspective in organisational frameworks of evaluation use.

Keywords

agency theory, evaluation use, legitimacy, organisational institutionalism, organisational theory, organisations

Corresponding author:

Tomasz Kupiec, Center for European Regional and Local Studies (EUROREG), University of Warsaw, ul. Krakowskie Przedmieście 30, 00-927 Warsaw, Poland.

Email: tomasz.kupiec@uw.edu.pl

Introduction

Use is a central concept in the theory, and practice of evaluation. This fact is evinced in the prominence of use in the field's professionalising documents (evaluation standards and guiding principles of the prominent evaluation societies and organisations), its centrality in various evaluation theories, as well as researchers' persistent interest in the subject (King and Alkin, 2019).

For several decades, evaluators and evaluation scholars have discussed and investigated the insufficient use of evaluation in decision-making (e.g. Alkin et al., 1979; Knorr, 1977; Weiss, 1972; Weiss and Bucuvalas, 1980). Several evaluation theories are specifically dedicated to use (especially Utilisation-Focused Evaluation, but also Participatory, Developmental, or Empowerment Evaluation) (Christie and Alkin, 2013). The rich body of evaluation literature contains numerous lists and frameworks with factors influencing evaluation use, sometimes comprising more than 50 elements (Huberman and Gather Thurler, 1991; Leviton and Hughes, 1981). Popular literature reviews (Cousins and Leithwood, 1986; Johnson et al., 2009) list over 100 empirical studies concerning evaluation use conducted between 1971 and 2005, with many more in the following years (e.g. Kupiec, 2015; Pattyn and Bouterse, 2020).

While these studies have strongly advanced the field, organisational theory lenses on evaluation use are still relatively underdeveloped. As argued by Raimondo (2018: 26), 'Organisational theories are conspicuously absent from the existing sense-making frameworks that evaluation scholars have put together to understand the use and influence of evaluation'. With evaluations being predominantly conducted by and for organisations and evaluation 'systems' becoming more normalised (Leeuw and Furubo, 2008), a comprehensive understanding of how organisational determinants influence evaluation use is indeed quintessential. It has also been argued that an organisational perspective might be the answer to the disturbing gap in the field, that is, the inability to plausibly explain the prevalence of symbolic use or non-use of evaluation (Højlund, 2014).

The few studies that do incorporate organisational theories adopt varying approaches. The traditional rational-objectivist model of evaluation (Schwandt, 1997; Van der Knaap, 1995) is closely related to rational choice theory and assumes that evaluation studies support reasoned choices. More recent studies have applied other organisational theories, most notably organisational institutionalism, according to which evaluation is meant to secure legitimacy from external actors (Ahonen, 2015), but also agency theory, which emphasises that evaluation is conducted to satisfy information needs of funders and supervising bodies (Weaver, 2007), and resource dependence theory, in which evaluation is considered as imposed on an organisation by the entities it is dependent upon (Eckerd and Moulton, 2011). Based on these theories, authors such as Eckerd and Moulton (2011), Højlund (2014) and Raimondo (2018) have proposed frameworks to explain evaluation use in terms of factors related to the context the organisation operates in and its role in the environment. Their efforts are very promising and have the potential to explain different functions and types of use (including symbolic, legitimising, or substantiating uses), which are difficult to comprehend under a mere rationalist lens.

However, the existing literature suffers from several deficiencies. The theoretical contributions of different authors have sometimes led to divergent predictions. Moreover, and more concerning, theoretical propositions have not always been in line with empirical accounts. These observations call for an in-depth reflection on how organisational literature on evaluation use can be further developed, a call to which this article responds.

The goal of this article is to advance knowledge on the determinants of evaluation use from an organisational theory perspective. At its core is a discussion of existing frameworks concerning the organisational determinants of evaluation use. We analyse key differences between them and examine how these frameworks do not always align with empirical observations. Furthermore, this article discusses the merits of two potential areas for future research on the organisational determinants of evaluation use, which can help align the theoretical contributions regarding evaluation use with empirical reality. First, we explain the relevance of considering organisational legitimacy as a potential direct determinant of evaluation use. We argue that the way an organisation seeks to justify its actions and existence in a social context (Suchman, 1995) may explain how it utilises evaluation. Second, we highlight the need to approach evaluation use from a more dynamic perspective. As evaluation practice and use may evolve over time, research may well consider these different stages of development when investigating determinants for use. Together, both these research avenues can enhance organisational studies of evaluation use.

Overview of the existing literature on the organisational determinants of evaluation use

The background – organisational theories

As stated in the introduction, the traditional rational-objectivist model of evaluation (Schwandt, 1997; Van der Knaap, 1995), which treats organisations as rational, goal-oriented entities seeking to maximise their efficiency (Sanderson, 2000), is closely related to rational choice theory despite not being explicitly stated as such in the literature. The core idea of this theory is that organisations and individuals will make reasoned choices about the desirability of adopting different courses of action (Dunn, 1981). These tenets correspond well with the often-stated primary goal of evaluation, which is to provide information that supports decision-making about programmes, policies and other actions (e.g. Fleischer and Christie, 2009; House, 1980; Patton, 1997). This basic and perhaps most intuitive type of evaluation use – known as instrumental use – was the first that scholars and practitioners tried to capture and measure. As it turned out, however, the direct use of evaluation findings was, overall, very limited (Weiss, 1998). One of the early outcomes of this observation was the idea to distinguish between knowledge for action (instrumental use) and knowledge for understanding (conceptual use), that is, knowledge influencing the minds of decision-makers without an immediate impact on their decisions (Rich, 1977). The consideration of conceptual use beyond instrumental use allowed for a more comprehensive account of the influence of evaluation on decision-makers.

Other types of evaluation use, such as symbolic and legitimising uses, were also recognised relatively early (Pelz, 1978), although some have claimed (Højlund, 2014) that they have never been appropriately integrated into the rational, ‘causal logic of evaluation use’ models, which expect an evaluation to lead directly to social betterment. In the words of Kirkhart (2000: 6), these types of use are secondary, ‘tacked on’, and support result-based use (instrumental and conceptual). Therefore, while rationalist evaluation theory recognises other types of evaluation use, it does not convincingly explain why they are so prevalent. For such explanations and justifications, scholars have turned to other organisational theories. For example, Albæk (1996) argued that organisations seen as political and cultural systems are expected to use evaluation to maximise their goals beyond instrumental ways. In a novel contribution, Carman (2011) proposed a theoretical framework for understanding the motivations behind

nonprofit evaluation that incorporates agency theory, institutional theory, resource dependence theory and stewardship theory. In over three-quarters of the cases studied, the first three theories turned out to offer the most accurate description of why organisations engaged in evaluation and how they used it.

Overall, institutional theory (particularly organisational institutionalism) and agency theory have received the most attention in the scholarly literature on evaluation use. *Agency theory* deals with task delegation from a principal to an agent in a situation where parties have divergent interests (Linder and Foss, 2015). Among other issues, it spells out how goal conflict, information asymmetry and outcome uncertainty can arise in principal–agent relations (Eisenhardt, 1989). Evaluation can help mitigate these problems by addressing the information needs of funders, supervising bodies and commissioners, ensuring that agents deliver on the objectives of principals (e.g. Carden, 2013; Weaver, 2007). In many instances, agents engage in evaluation because funders require them to do so (Carman, 2011). This function of evaluation is prevalent, especially in large (including international) organisations or multi-organisational settings (Eckhard and Jankauskas, 2020; Picciotto, 2016).

Organisational institutionalism considers how the behaviour of organisations is influenced by the rules, norms and ideologies of society (Meyer and Rowan, 1983) or a common understanding of what is appropriate (Zucker, 1983). Organisations conform to these rationalised myths to appear rational and gain legitimacy (Scott, 1983). In this line of thinking, the expected initial goal of evaluation is to secure legitimacy from external actors by signalling a commitment to knowledge utilisation (Ahonen, 2015). Therefore, the force that shapes evaluation practice is environmental pressure (Carman, 2011). Evaluation from this perspective is an organisational recipe, a generally accepted ritual legitimised by expectations in organisational environments (Dahler-Larsen, 2012).

Somewhat similar conclusions about the motivation to conduct an evaluation arise from *resource dependence theory*, which assumes that organisations are embedded in their environment and their actions are a function of the environment because organisations must acquire resources from external sources to survive (Pfeffer and Salancik, 2003). According to this theory, organisations use evaluation to promote their activities and get something in return (Behn, 2003), but the resource providers (and their expectations) may, to some extent, control or shape evaluation activities (Eckerd and Moulton, 2011).

These three theoretical perspectives are not fully distinguishable; rather, they overlap and complement one another because they all perceive organisations as open systems embedded in and dependent on their broader environment and acquiring resources while simultaneously gaining social support and legitimacy from it (Scott and Davis, 2015). As will be presented below, existing evaluation use frameworks often combine insights from two or all three of these theories. This is in line with the general tendency in organisational studies to combine several theories, for instance, institutional theory with agency theory and resource dependence theory (Greenwood et al., 2017).

Frameworks for the organisational determinants of evaluation use

With these organisational theories in mind, we turn to discussing the existing frameworks of evaluation use. At their basis, each deals with one or several types of evaluation use as dependent variables, which they explain by relying on variables derived from the above-mentioned organisational theories perceiving organisations as open systems. Building on the work of Johnson (1998), we have also included so-called ‘implicit’ frameworks, that is, contributions where the variables and/or relationships between them are not directly depicted but can be logically implied.

The literature presented in this section was not identified through a rigorous systematic review. We cannot, therefore, attest to its exhaustiveness; however, we believe that our search strategy allowed us to capture the important contributions and full diversity of approaches to explain evaluation use from the lens of open system organisational theories.¹ As mentioned, theoretical studies seeking to understand evaluation use from an organisational perspective are relatively scarce (Raimondo, 2018).

An interesting contribution explicitly based on organisational institutionalism was proposed by Højlund (2014). His framework predicts variation in types of evaluation use resulting from two variables: external pressures for adopting evaluation, such as regulations, cultural constraints, uncertainty, or normative expectations from the environment, and an internal propensity to evaluate. Regarding the latter, two extreme positions were distinguished based on an organisation's characteristics (borrowed from Brunsson, 2002): the 'political' organisation, which derives legitimacy from talks and decisions, and the 'action' organisation, which depends on its ability to produce outputs. The combination of these factors produces four possible *adoption modes* of evaluation practice: coercive, mimetic, normative and voluntary. These modes can be read as a clear reference to the mechanisms of institutional isomorphism conceptualised by DiMaggio and Powell (1983), with each leading to a different type of evaluation use (or group of types).

External pressure, or the external environment, is also included in the framework by Eckerd and Moulton (2011), where it is the primary determinant of the evaluation method.² They argued that the type of evaluation use is contingent upon the role of the organisation in society. This also suggests that evaluation systems might be subject to loose coupling, that is, there may be discrepancies between expected, declared and actual evaluation use.

Decoupling – a central concept in organisational institutionalism (Boxenbaum and Jonsson, 2017) – received particular attention in the study by Raimondo (2018). She argued that while evaluation systems are rooted in a willingness to remedy organisational loose coupling, they may elicit patterns of behaviours that contribute to further organisational decoupling (Raimondo and Leeuw, 2022). In her framework of evaluation use, the way organisations adopt and use evaluation is a result of the interplay between cultural (e.g. conflicting norms across internal units or competing definitions of success among stakeholders) and material forces (e.g. available resources, the relative power of donors and clients) inside and outside of the organisation. Both cultural and material factors have internal and external origins, which suggests that open system organisational theories are not sufficient in explaining evaluation use.

Although informed by open system organisational theories, Carman (2005) and Lall (2015) adopted somewhat different approaches. Carman conceived each theory as a specific motivation that may lead to evaluation. For instance, in the spirit of agency theory, evaluation is required by funders, the board, or management. For resource dependence theory, evaluation helps secure resources and promote the organisation to stakeholders. Lall translated the same theories into the relationships with funders and other stakeholders. For instance, whereas agency theory is particularly apt to describe situations in which an organisation previously received grant funding, resource dependence theory lends itself well to situations in which organisations seek grant funding. As such, in their empirical studies, an organisation's motivation to evaluate or its relationship with funders and other stakeholders influences the type of evaluation use.

Table 1 summarises the main tenets of the frameworks discussed, including how they connect organisational theories with types of evaluation use.

In addition to these five frameworks, several other contributions are worth mentioning. While they do not directly engage with open system organisational theories to explain

Table 1. Evaluation use frameworks.

| Use factors (independent variables) | Use types (dependent variables) | Type(s) of organisations | Underlying theories | Author(s) |
|---|---|------------------------------------|---|---------------------------|
| External pressure, internal propensity to evaluate | Instrumental, conceptual, legitimising, symbolic, non-use | NGOs, various public organisations | Organisational institutionalism | Højlund (2014) |
| External environment, strategic role of the organisation | Instrumental, conceptual, symbolic | NGOs | Organisational institutionalism, contingency, resource dependence | Eckerd and Moulton (2011) |
| External material factors, external cultural factors ^a | Learning (<i>instrumental</i> ^b , <i>conceptual</i>), accountability (<i>symbolic</i> , <i>substantiating</i>) | International organisations | Organisational institutionalism, principal-agent (among others) ^c | Raimondo (2018) |
| Relationship with funders (previously received grant or seeking grant, receiving support from incubators), role of an organisation (social vs financial objectives) | Funder-centric (measure to prove), other (measure to improve) | Social enterprises | Organisational institutionalism, principal-agent, stewardship | Lall (2015) |
| Motivations to evaluate: requirement from funders, the board, management, securing resources, promoting organisation to funders, part of everyday management practice | <i>Instrumental</i> , <i>conceptual</i> , <i>legitimising</i> , <i>symbolic</i> ^b | NGOs | Organisational institutionalism, principal-agent, resource dependence (among others) ^c | Carman (2005) |

^aThis framework combines external and internal dimensions. In the table, only external factors are depicted.

^bElements that are not explicitly indicated in the framework but can be implied are presented in *italic*.

^cAlternatives to open system organisational theories were also referred to.

evaluation use, they approach the subject from an organisational lens and refer to the factors that can be found in the frameworks mentioned. For example, Bryan et al. (2021) suggested that the way evaluation is used, or specifically, the capacity to use evaluation in a particular way, depends on the dominant type of accountability. They demonstrate that lateral (e.g. to volunteers) and downwards (e.g. to participants) accountabilities are more likely to support learning from evaluation than upwards accountability (e.g. to regulators). Picciotto (2016) provided a noteworthy example of approaching evaluation use from the lens of neo-institutional economics and agency dilemmas. Building on the work of De Laat (2013), he described how different evaluation governance designs steer the evaluation function. Evaluation governance designs, in his understanding, differ depending on whether there is a conflation or clear separation of the roles of the commissioning body, evaluator, decision maker, and beneficiary. In what he termed the ‘market-oriented configuration’, where decision-makers commission an independent evaluator, learning is generally expected to be limited. In turn, a so-called self-evaluation configuration, where the commissioner, evaluator and decision maker are the same entity, will foster an instrumental use of evaluation findings. Only an evaluation governance design keeping the commissioner, evaluator, and decision maker at arm’s length will be conducive to both accountability and learning. In this context, it is also relevant to mention the work of Eckhard and Jankauskas (2020), who studied the distribution of control over evaluation resources between member states and the administration of international organisations. This distribution turns out to be influential in determining an evaluation’s focus on accountability or learning. As such, their study empirically confirms the relevance of external material factors (see Raimondo, 2018).

Challenges for future research

Differences between frameworks

The frameworks mentioned above clearly share some key characteristics, which is a logical reflection of the fact that they are anchored in the same organisational theories. These commonalities relate, for instance, to the assumption that evaluation use is determined by the impact of external pressure and an emphasis on the role of an organisation in society.

However, when comparing these frameworks, more differences than commonalities stand out, and they do not result just from the fact that some frameworks are grounded in empirical studies (Carman, 2005; Lall, 2015) while others are more theoretically focused (Højlund, 2014; Raimondo, 2018).

First, the attention of the contributions differs, with frameworks focusing on different types of organisations, such as NGOs (Carman, 2005; Eckerd and Moulton, 2011), large international organisations (Raimondo, 2018) or social enterprises (Lall, 2015). Only Højlund (2014) refers to various organisations (small NGOs, universities, government agencies, and municipalities), but he only uses them as examples to illustrate ideal adoption modes.

Second, while each contribution relies on the same set of organisational theories, they engage with them differently. In some contributions, organisational theories constitute a basis for identifying detailed and numerous determinants of use, such as the intensity of external pressures (Eckerd and Moulton, 2011; Højlund, 2014) or a longer list of material and cultural factors (Raimondo, 2018). For Carman (2005), each theory leads to a distinct motivation for evaluation. Lall (2015), however, operationalises each theory in terms of the relationship between an organisation and its stakeholders.

Third, similar concepts arising in the frameworks are approached differently in practice. A prime example is the role of the organisation. Højlund (2014) distinguished between action and political organisations, Eckerd and Moulton (2011) proposed six more detailed roles for NGOs (e.g. service provision or social capital creation) and Lall (2015) considered the extent to which organisations prioritise social versus financial objectives.

Finally, each framework relies on different typologies of evaluation use or related proxies. Eckerd and Moulton (2011) differentiated between instrumental, conceptual, and symbolic use. Højlund (2014) included legitimising use and non-use as separate categories. However, what he considered as legitimising use was termed as conceptual (satisfying funders' requirements) or symbolic use (demonstrating legitimacy) by Eckerd and Moulton (2011). Raimondo (2018), on her part, referred to what is generally understood as evaluation functions, that is, accountability and learning, which only loosely relate to use types (learning corresponding more with instrumental and conceptual use, and accountability with symbolic and legitimising use). Carman (2005) formulated ten statements to operationalise evaluation use (e.g. to develop new programmes or report to funders), which are all connected to the common types of use (instrumental, conceptual, symbolic, and legitimising), though not always in an unambiguous way.

These variations complicate direct comparisons between the frameworks. Moreover, the frameworks also differ in how they present the role of specific variables and the relationship between them. For instance, Højlund (2014) suggested that the type of evaluation use is determined by external pressures and the role of the organisation. Conversely, Eckerd and Moulton (2011) argued that the external environment may influence the evaluation's scope, but the type of evaluation use will ultimately depend on the role that the organisation serves.

Nuanced differences also exist when it comes to the motivation for adopting evaluation. Raimondo (2018) claimed that evaluation is often initiated to secure the interest of the principal and that the first and foremost function of an evaluation is to legitimise the agent organisation (associated with symbolic or legitimising use). Højlund (2014), Eckerd and Moulton (2011), and Carman (2005), however, also highlighted the possibility of an evaluation being conducted to serve learning purposes (instrumental or conceptual use). Again, while the same variables feature in the different frameworks, their role in each assumes subtle differences.

Empirical observations not fully explained by current frameworks

Evaluation literature offers many empirical observations of organisations using evaluation, that may serve as a verification basis for the theoretical propositions of the frameworks mentioned above. It appears that there are at least three types of situations where predictive capabilities of the frameworks are lacking: (1) when organisations coerced to conduct evaluation, move beyond symbolic use and use evaluation's findings instrumentally or conceptually; (2) when organisations operate in the same context and face similar external pressures, but use evaluation differently; and (3) when organisations operate under seemingly stable external pressure, but the way they use evaluation evolves. We elaborate on each of these instances in more detail below.

First, evaluation triggered by external pressure rather than internal propensity do not necessarily result in the non-use or symbolic use of the evaluation. This was observed in the case of Australian (Kelly, 2021) and American NGOs (Carman, 2005). In fact, in some contexts, increased external pressure can facilitate instrumental and conceptual use. Research comparing Polish and Spanish regional authorities evaluating EU cohesion policy programmes, for

example, confirmed this. In Poland, the influence of the external environment was not only limited to coercing evaluation but also contributed to evaluation capacity building (Wojtowicz and Kupiec, 2018). Similar mechanisms can be found in the works of Anderson (2013), Pattyn (2015) and Torres and Preskill (2001).

Second, differences in the types and levels of evaluation use may occur among organisations or units operating under similar external pressures. This was observed in the Directorates-General and Services of the European Commission (Williams et al., 2002). Although they were subjected to the same ‘soft’ standards and guidelines, strong variations in levels of use were found. Another example was illustrated by Borum and Hansen (2000), who reported differences in the use of evaluation results in three departments of the same higher education institution in Denmark and, therefore, largely operating under the same organisational framework. Nonetheless, differences in departmental characteristics, such as their structure and history, were said to explain the variations in their evaluation use.

Third, the literature provides examples of organisations operating under relatively constant external pressure yet evolving in the way they use evaluation. For example, Preskill and Boyle (2008) described cases where the attitudes of NGOs towards evaluation changed over time. While evaluation was initially conceived as a punitive measure conducted to satisfy funders’ requirements, it became gradually regarded as a tool to support decisions and increase efficiency, and as such also catering to internal demands. In their study of community-based organisations conducting HIV prevention programmes, Gibbs et al. (2002) noted a similar shift, wherein evaluation practice evolved from simple compliance with funding requirements to more intrinsically motivated evaluation studies set up to improve interventions and gain a broader understanding of the mechanism behind them. They argued that this evolution was triggered by internal resources, such as staff and leadership. Internal factors were also found to be decisive in how evaluation use was conceived in Kelly’s (2021) study of Australian NGOs.

Examples of how evaluation use can evolve over time are found in the public sector as well. For instance, House et al. (1996) examined how the Education and Human Resources Directorate of the US National Science Foundation was obliged to conduct evaluation by the US Congress in the 1990s. In the beginning, these studies were merely aimed at ensuring accountability due to a lack of awareness of other possible functions of evaluation. With time and growing experience, the organisation built an evaluation culture that became an integral part of its management practice.

Directions for future research

The analysis above discussed the merits of existing organisational frameworks of evaluation use while also examining their variances and limitations in explaining empirical observations. Building upon this analysis, this section suggests two research avenues that are worth exploring in more depth, given their potential to address the above-mentioned challenges. The first avenue proposes to consider organisational legitimacy more explicitly as a determinant of evaluation use. The second advocates including the element of dynamics in organisational frameworks of evaluation use.

Organisational legitimacy as a determinant of evaluation use

Organisational legitimacy is a central concept in organisational research and features prominently in open system organisational theories (Deephouse and Suchman, 2008). It can be perceived as

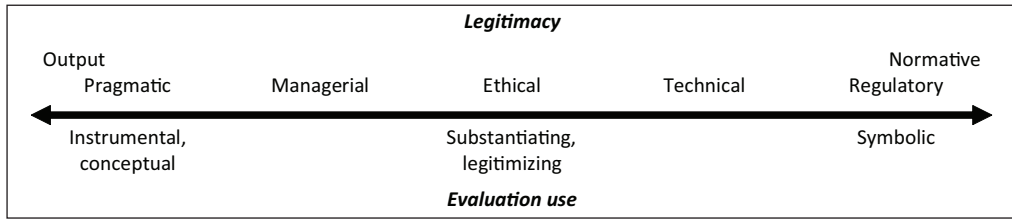


Figure 1. An initial assumption on the possible relationship between legitimacy type and evaluation use type.

Source. Own elaboration based on Díez-de-Castro et al. (2018) and Meyer and Rowan (1977).

both a dynamic constraint (Dowling and Pfeffer, 1975) or a tool to acquire more resources (Meyer and Rowan, 1977; Pfeffer and Salancik, 2003).

While legitimacy can be found as an underlying or related element in the frameworks discussed above (demonstrated in Table 1), no one has explicitly treated it as a determinant of evaluation use. Nevertheless, organisational legitimacy may be an important influencing factor to consider. In previous studies, for instance, the sources of organisational legitimacy proved to influence organisational routines (Johansson and Sell, 2004). Furthermore, the type of organisational legitimacy itself is determined by a range of the organisation's characteristics (Dowling and Pfeffer, 1975). From this, it follows that the dominant type, source or dimension of legitimacy may be a more general and universal determinant of evaluation use than how it has been presented in the existing frameworks.

To explore the role of organisational legitimacy in a more systematic way, inspiration can be drawn from previous studies on legitimacy outside the field of evaluation. For example, Díez-de-Castro et al. (2018) developed a detailed typology distinguishing between regulatory, technical, ethical, managerial, and pragmatic legitimacies that lends itself well to studying the role of legitimacy in influencing particular types of evaluation use. For instance, it is plausible that organisations seeking regulatory legitimacy derived from compliance with rules and norms will use evaluation symbolically to demonstrate this compliance. Relatedly, organisations resting on moral or ethical legitimacy, supporting principles which are perceived as socially beneficial, may be more inclined to seek substantiating or legitimising evaluation uses. Organisations whose legitimacy depends on securing stakeholders' interests, defined as pragmatic legitimacy, can be expected to use evaluation more instrumentally.

In the same vein, it is worth reflecting on how evaluation use may differ between production and institutionalised organisations, in line with the classic typology described by Meyer and Rowan (1977). Production organisations are typically characterised by output legitimacy and strong output controls, while institutionalised organisations rely more on normative legitimacy, in which confidence is achieved through compliance with institutional rules.

A combination of both typologies can constitute a skeleton for more systematic research on the topic. Figure 1 presents how different types of legitimacy may manifest in various types of organisations and how this can affect evaluation use. Future research would ideally test these assumptions in a systematic way. Figure 1 is relevant for a multitude of organisations ranging from small NGOs to large international organisations.

One may also incorporate other typologies of legitimacy to fine-tune and test insights about evaluation use. Here, the seminal distinction between input and output legitimacy is relevant

to consider (Scharpf, 1999). Whereas output legitimacy is grounded in policy processes making use of expert debates and findings, input legitimacy emphasises strategic bargaining and coalition building (Skogstad, 2003). In the field of evaluation, rational-objectivist approaches tend to prioritise output legitimacy, while argumentative-subjectivist evaluation approaches predominantly emphasise input legitimacy. One can hypothesise that instrumental use and conceptual use are more at stake when organisations prioritise output legitimacy, whereas symbolic and substantiating uses are likely more associated with organisations seeking input legitimacy.

Accounting for a possible evolution of evaluation use

In addition to unravelling what different types of legitimacy mean for evaluation use, the field is also in need of approaching evaluation use in a more dynamic way. The frameworks discussed are primarily static (Martinaitis et al., 2018), suggesting that initial conditions determine how an organisation will use evaluation in the future and disregarding changes in evaluation practice. While Raimondo (2018) argued that it is very difficult to reorient an established evaluation system, other authors (e.g. Højlund, 2014) do not refer to potential changes in evaluation practice and use at all.

The static representation of evaluation is not fully in line with actual empirical findings. Several examples contradicting this representation have been mentioned in Section 'Empirical observations not fully explained by current frameworks'. In addition, one can find evidence of a dynamic changes of evaluation practice in multi-organisational policy evaluation systems in EU countries (Kupiec et al., 2020), developing countries (Horton, 1999), European Commission Directorates (De Francesco, 2019), government agencies (House et al., 1996; Pattyn, 2014) or NGOs (Love, 1998). These observations indicate that how evaluation is conducted and used may or even is expected to change over time. We argue that much of the discrepancy between empirical observations and existing frameworks, as well as the inconsistencies between frameworks, can be explained by accounting for time and dynamism. In other words, a holistic framework of evaluation use should recognise that there are several stages in the development of evaluation practice.

To our knowledge, there are no studies conceptualising the stages in the evolution of evaluation use. However, valuable inspiration can be found in the literature on evaluation capacity building. For example, Bourgeois and Cousins (2008) identified four stages through which evaluation capacity building typically proceeds: (1) traditional evaluation – with externally coerced evaluation activities; (2) awareness and experimentation – when the organisation learns about the benefits of evaluation; (3) implementation – where evaluation becomes more clearly defined in the organisation; and (4) adoption – when evaluation becomes regular and sufficient financial and human resources are allocated to it.

A related, albeit slightly different, perspective is presented by Gibbs et al. (2002) (see also Gilliam et al., 2003), who argued that evaluation capacity may develop in three stages: (1) compliance – when an evaluation is conducted to comply with external forces without apparent benefits, (2) investment – allocating internal resources to evaluation may improve interventions and support funding expansion, (3) advancement – more ambitious evaluations are conducted which can contribute to a broader understanding of theory and practice.

Both proposals have much in common, but the distinction between the stages is not always clear, and it may well be the case that organisations do not move through every stage. For this

reason, and considering the relatively premature state of organisational research on evaluation use, we believe that much can be gained by rigorously distinguishing between two core stages of evaluation development.

- *Adoption phase* – evaluation practice is adopted in response to external pressure (directly exerted by other organisations or perceived indirectly and related to uncertainties in the field).
- *Advancement phase* – evaluation is carried out in the organisation for some time, long enough to build capacity to conduct it and to reflect on its possible role.

To further the field of evaluation, it is essential to gain a more systematic understanding of the conditions and mechanisms triggering organisations to move from one stage to another. We believe that the context of an organisation can be decisive in this regard, including the nature and structure of the tasks it oversees, the intervention cycles the organisation deals with, or the funding cycle it operates in (Devine, 2002). Evaluation awareness and attitudes towards evaluation may also change simply as a result of exposure to evaluation practice, which may yield a more ‘organic process’ (Preskill and Boyle, 2008). Therefore, research on evaluation use may benefit from complementing an open system perspective with a more natural system approach, that is, combining an external perspective with an internal one. Such an approach has been adopted in the related field of expert knowledge use. For example, Rimkutė (2015) combined external factors, including formal and informal external pressure, with internal factors, such as the capacity to produce scientific outputs.³ Schrefler (2010) proposed two external explanatory factors – level of conflict in the policy arena and level of problem tractability – but suggested that the capacity of an agency should be used as an additional control variable which can influence the use of knowledge.

Findings from the field of expert knowledge use imply that the capacity to conduct evaluation may be another crucial factor influencing evaluation use. Our additional hypothesis and suggestion for future research is that evaluation capacity may be a force pushing organisations from the adoption to the advancement stage of evaluation practice, and that different factors can determine evaluation use at either stage. While the external pressure to evaluate may be decisive at the adoption phase, the role of the organisation or its type of legitimacy may be more impactful at the advancement phase.

Including evaluation capacity in the equation may also lead to the optimistic conclusion for evaluation practitioners that organisations failing to use evaluation or using it in an undesired way are not necessarily pre-determined to do so in the future. Organisations may learn to use evaluation as a consequence of involvement in the evaluation process (Patton, 1997) or may benefit from so-called *process use* (as opposed to *findings use*). To the extent possible, evaluators may also help organisations build evaluation capacity (Cousins et al., 2004) to facilitate moving from the adoption to the advancement phase.

Conclusion

The aim of this article was to further the understanding of organisational factors of evaluation use, that is determinants of evaluation use grounded in organisational theories. This approach resonates with the increasing attention on evaluation systems, in which emphasis is put on

streams of studies produced in a certain epistemic context (Rist and Stame, 2006) in line with sociological system thinking (Leeuw and Furubo, 2008). Scholars advocating to approach evaluation from a systemic lens have suggested shifting focus away from an 'evaluation-centric' perspective with an evaluation study as a unit of analysis (Højlund, 2014) and relocating organisational and institutional factors from the periphery to the centre of theoretical frameworks (Dahler-Larsen, 2012). Our approach corresponds with this emerging field of research and can provide a more comprehensive understanding of evaluation use.

As it currently stands, contributions explaining evaluation use from the lens of organisational theories are still scarce. Those available have usually adopted an open system perspective and relied on organisational institutionalism, agency theory, or resource dependence theory.

Our analysis focused on five frameworks centred around the relationship between types of evaluation use and explanatory variables derived from open system organisational theories. While anchored in similar theoretical principles, the frameworks demonstrate more differences than commonalities by emphasising different types of organisations, for instance, and relying on different typologies of evaluation use, which limits the possibilities for comparisons. Furthermore, the frameworks differ in terms of the role attributed to specific factors, such as the role of external pressure in triggering evaluation practice and use.

These differences alone are curious and necessitate further research on the issue. More importantly, these existing frameworks fall short of explaining key empirical observations. The available literature does not provide adequate explanations for why a dominant type of evaluation use can shift over time or differ across organisations operating under seemingly stable external conditions.

Responding to these challenges, we offer two directions to advance theoretical and empirical research on the topic. The first maintains an open system organisational theory perspective but revolves around organisational legitimacy as a potential determinant of evaluation use. As we explained, organisational legitimacy may be a more decisive factor than how it is currently conceived in existing frameworks, and a systematic study of the role and different types of organisational legitimacy may help explain evaluation use in a broader range of organisations in both the public and non-government sectors. The second approach incorporates a natural system perspective and suggests a more rigorous treatment of time and dynamics in theorising evaluation use. Distinguishing between the several stages in the development of evaluation practice can help capture why evaluation use evolves and why organisations operating with similar external pressures may use evaluation differently.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: The work of the T.K. and D.C.-J. was supported by the National Science Centre, Poland, grant number 2019/33/B/HS5/01336. V.P. acknowledges the support by the Dutch Research Council (Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) (grant no VI.Veni.211R.060)

ORCID iDs

Tomasz Kupiec  <https://orcid.org/0000-0002-6469-7746>

Dorota Celińska-Janowicz  <https://orcid.org/0000-0002-0101-7475>

Valérie Pattyn  <https://orcid.org/0000-0001-7844-8858>

Notes

1. Our review strategy combined a SCOPUS database search of records referring to institutionalism, agency theory, or resource dependence and evaluation use or evaluation utilisation with back and forward reference and citation tracking for key contributions. Earlier studies suggested that snowballing is a more effective strategy to identify relevant contributions than predefined, protocol-driven database searches (Greenhalgh and Peacock, 2005). In our case, this was also proven true. On many occasions, authors refer to a particular organisational theory only in the main body of the text and not in the abstract, title, or keywords.
2. While the authors use the term *evaluation method*, one could argue that what they refer to is rather the evaluand, that is, the evaluation focus.
3. Her dependent variables included problem-solving use (which can be related to instrumental use or conceptual use), strategic use, and symbolic use.

References

- Ahonen P (2015) Aspects of the institutionalization of evaluation in Finland: Basic, agency, process and change. *Evaluation* 21(3): 308–24.
- Albæk E (1996) Why all this evaluation? Theoretical notes and empirical observations on the functions and growth of evaluation, with Denmark as an illustrative case. *Canadian Journal of Program Evaluation* 11(2): 1–34.
- Alkin MC, Daillak R and White P (1979) *Using Evaluations: Does Evaluation Make a Difference?* Beverly Hills, CA: Sage.
- Anderson KL (2013) *A Retrospective Case Study of Long-Term Evaluation Capacity Building at Neighborhood House*. PhD Thesis, University of Minnesota, Minneapolis, MN.
- Behn RD (2003) Why measure performance? Different purposes require different measures. *Public Administration Review* 63(5): 586–606.
- Borum F and Hansen HF (2000) The local construction and enactment of standards for research evaluation: The case of the Copenhagen business school. *Evaluation* 6(3): 281–99.
- Bourgeois I and Cousins JB (2008) Informing evaluation capacity building through profiling organizational capacity for evaluation: An empirical examination of four Canadian federal government organizations. *Canadian Journal of Program Evaluation* 23(3): 127–46.
- Boxenbaum E and Jonsson S (2017) Isomorphism, diffusion and decoupling: Concept evolution and theoretical challenges. In: Greenwood R, Oliver C, Lawrence TB, et al. (eds) *The SAGE Handbook of Organizational Institutionalism*. London: Sage, 79–104.
- Brunsson N (2002) *The Organization of Hypocrisy: Talk, Decisions and Actions in Organizations*. Copenhagen: Abstrakt & Liber.
- Bryan TK, Robichau RW and L'Esperance GE (2021) Conducting and utilizing evaluation for multiple accountabilities: A study of nonprofit evaluation capacities. *Nonprofit Management and Leadership* 31(3): 547–69.
- Carden F (2013) Evaluation, not development evaluation. *American Journal of Evaluation* 34(4): 576–9.
- Carman JG (2005) *Program Evaluation Use and Practice: A Study of Nonprofit Organizations in New York State*. Albany, NY: State University of New York.
- Carman JG (2011) Understanding evaluation in nonprofit organizations. *Public Performance & Management Review* 34(3): 350–77.

- Christie CA and Alkin MC (2013) An evaluation theory tree. In: Alkin MC (ed.) *Evaluation Roots: Tracing Theorists' Views and Influences*. Thousand Oaks, CA: Sage, 11–57.
- Cousins JB and Leithwood KA (1986) Current empirical research on evaluation utilization. *Review of Educational Research* 56(3): 331–64.
- Cousins JB, Goh SC, Clark S, et al. (2004) Integrating evaluation inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation* 19(2): 99–141.
- Dahler-Larsen P (2012) *The Evaluation Society*. Stanford, CA: Stanford University Press.
- De Francesco F (2019) Following the evidence practice: An analysis of evaluation studies on EU railway policy. *Politische Vierteljahresschrift* 60(4): 785–803.
- De Laat B (2013) The tricky triangle: Evaluator, evaluand, evaluation commissioner. In: Laubli Lound M and Mayne J (eds) *Enhancing Evaluation Use: Insights from Internal Evaluation Units*. Thousand Oaks, CA: Sage, 15–36.
- Deephouse DL and Suchman M (2008) Legitimacy in organizational institutionalism. In: Greenwood R, Oliver C, Sahlin K, et al. (eds) *The SAGE Handbook of Organizational Institutionalism*. London: Sage, 49–77.
- Devine DJ (2002) A review and integration of classification systems relevant to teams in organizations. *Group Dynamics: Theory, Research and Practice* 6(4): 291–310.
- Díez-de-Castro E, Peris-Ortiz M and Díez-Martín F (2018) Criteria for evaluating the organizational legitimacy: A typology for legitimacy jungle. In: Díez-de-Castro E and Peris-Ortiz M (eds) *Organizational Legitimacy*. Cham: Springer, 1–21.
- DiMaggio PJ and Powell WW (1983) The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review* 48(2): 147–60.
- Dowling J and Pfeffer J (1975) Organizational legitimacy: Social values and organizational behavior. *The Pacific Sociological Review* 18(1): 122–36.
- Dunn WN (1981) *Public Policy Analysis: An Introduction*. Englewood Cliffs, NJ: Prentice Hall.
- Eckerd A and Moulton S (2011) Heterogeneous roles and heterogeneous practices: Understanding the adoption and uses of nonprofit performance evaluations. *American Journal of Evaluation* 32(1): 98–117.
- Eckhard S and Jankauskas V (2020) Explaining the political use of evaluation in international organizations. *Policy Sciences* 53(4): 667–95.
- Eisenhardt KM (1989) Agency theory: An assessment and review. *Academy of Management Review* 14(1): 57–74.
- Fleischer DN and Christie CA (2009) Evaluation use: Results from a survey of US American evaluation association members. *American Journal of Evaluation* 30(2): 158–75.
- Gibbs D, Napp D, Jolly D, et al. (2002) Increasing evaluation capacity within community-based HIV prevention programs. *Evaluation and Program Planning* 25(3): 261–9.
- Gilliam A, Barrington T, Davis D, et al. (2003) Building evaluation capacity for HIV prevention programs. *Evaluation and Program Planning* 26(2): 133–42.
- Greenhalgh T and Peacock R (2005) Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *The BMJ* 331(7524): 1064–5.
- Greenwood R, Oliver C, Lawrence TB, et al. (eds) (2017) *The SAGE Handbook of Organizational Institutionalism*. London: Sage.
- Højlund S (2014) Evaluation use in the organizational context—changing focus to improve theory. *Evaluation* 20(1): 26–43.
- Horton D (1999) Evaluation in developing countries: An introduction. *Knowledge, Technology & Policy* 11(4): 5–12.
- House ER (1980) *Evaluation with Validity*. Beverly Hill, CA: Sage.
- House ER, Haug C and Norris N (1996) Producing evaluations in a large bureaucracy. *Evaluation* 2(2): 135–50.

- Huberman M and Gather Thurler M (1991) *De la Recherche à la pratique: Eléments de base*. Berne: Peter Lang.
- Johansson AC and Sell J (2004) Sources of legitimation and their effects on group routines: A theoretical analysis. In: Jonson C (ed.) *Legitimacy Processes in Organizations*, vol. 22. Bingley: Emerald Group Publishing, 89–116.
- Johnson K, Greenesid LO, Toal SA, et al. (2009) Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation* 30(3): 377–410.
- Johnson RB (1998) Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning* 21(1): 93–110.
- Kelly LM (2021) *Evaluation in Small Development Non-Profits: Deadends, Victories and Alternative Routes*. Berlin: Springer Nature.
- King JA and Alkin MC (2019) The centrality of use: Theories of evaluation use and influence and thoughts on the first 50 years of use research. *American Journal of Evaluation* 40(3): 431–58.
- Kirkhart KE (2000) Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation* 2000(88): 5–23.
- Knorr KD (1977) Policymakers' use of social science knowledge: Symbolic or instrumental? In: Weiss CH (ed.) *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books, 165–82.
- Kupiec T (2015) Program evaluation use and its mechanisms: The case of cohesion policy in Polish regional administration. *Zarządzanie Publiczne* 3(33): 67–83.
- Kupiec T, Wojtowicz D and Olejniczak K (2020) Between compliance and systemic change – evaluation practice in eight CEE countries. In: Mazur S (ed.) *Public Administration in Central Europe. Ideas as Causes of Reforms*. London and New York: Routledge, 129–46.
- Lall SA (2015) *Measuring to Improve vs. Measuring to Prove: Understanding Evaluation and Performance Measurement in Social Enterprise*. PhD Thesis, The George Washington University, Washington, DC.
- Leeuw FL and Furubo JE (2008) Evaluation systems: What are they and why study them? *Evaluation* 14(2): 157–69.
- Leviton LC and Hughes EF (1981) Research on the utilization of evaluations: A review and synthesis. *Evaluation Review* 5(4): 525–48.
- Linder S and Foss NJ (2015) Agency theory. In: Wright J (ed.) *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam: Elsevier.
- Love AJ (1998) Internal evaluation: Integrating evaluation and social work practice. *Scandinavian Journal of Social Welfare* 7(2): 145–51.
- Martinaitis Ž, Christenko A and Kraučūnienė L (2019) Evaluation systems: How do they frame, generate and use evidence? *Evaluation* 25(1): 46–61.
- Meyer JW and Rowan B (1977) Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology* 83(2): 340–63.
- Meyer JW and Rowan B (1983) The structure of educational organizations. In: Meyer JW and Scott WR (eds) *Organizational Environments: Ritual and Rationality*. Beverly Hills, CA: Sage, 217–25.
- Patton MQ (1997) *Utilization-Focused Evaluation: The New Century Text*, 3rd edn. Thousand Oaks, CA: Sage.
- Pattyn V (2014) Why organizations (do not) evaluate? Explaining evaluation activity through the lens of configurational comparative methods. *Evaluation* 20(3): 348–67.
- Pattyn V (2015) Explaining variance in policy evaluation regularity. The Case of the Flemish public sector. *Public Management Review* 17(10): 1475–95.
- Pattyn V and Bouterse M (2020) Explaining use and non-use of policy evaluations in a mature evaluation setting. *Humanities and Social Sciences Communications* 7(1): 19.
- Pelz DC (1978) Some expanded perspectives on use of social science in public policy. In: Yinger JM and Cutler SJ (eds) *Major Social Issues: A Multidisciplinary View*. New York: Palgrave Macmillan, 346–57.

- Pfeffer J and Salancik GR (2003) *The External Control of Organizations: A Resource Dependence Perspective*. Stanford, CA: Stanford University Press.
- Picciozzo R (2016) Evaluation and bureaucracy: The tricky rectangle. *Evaluation* 22(4): 424–34.
- Preskill H and Boyle S (2008) Insights into evaluation capacity building: Motivations, strategies, outcomes and lessons learned. *Canadian Journal of Program Evaluation* 23(3): 147–74.
- Raimondo E (2018) The power and dysfunctions of evaluation systems in international organizations. *Evaluation* 24(1): 26–41.
- Raimondo E and Leeuw FL (2022) Evaluation systems and bureaucratic capture: Locked in the system and potential avenues for change. In: Perrin B and Tyrrell T (eds) *Comparative Policy Evaluation: Vol 29. Changing Bureaucracies: Adapting to Uncertainty and How Evaluation Can Help*. Abingdon: Routledge, 143–60.
- Rich RF (1977) Measuring knowledge utilization: Processes and outcomes. *Knowledge and Policy* 10: 11–24.
- Rimkutė D (2015) Explaining differences in scientific expertise use: The politics of pesticides. *Politics and Governance* 3(1): 114–27.
- Rist R and Stame N (eds) (2006) *From Studies to Streams: Managing Evaluative Systems*. London: Transaction Publishers.
- Sanderson I (2000) Evaluation in complex policy systems. *Evaluation* 6(4): 433–54.
- Scharpf F (1999) *Regieren in Europa: Effektiv und demokratisch?* Frankfurt and New York: Campus.
- Schrefler L (2010) The usage of scientific knowledge by independent regulatory agencies. *Governance* 23(2): 309–30.
- Schwandt TA (1997) Evaluation as practical hermeneutics. *Evaluation* 3(1): 69–83.
- Scott WR (1983) The organization of environments Network, cultural and historical elements. In: Meyer JW and Scott WR (eds) *Organizational Environments: Ritual and Rationality*. Beverly Hills, CA: Sage, 199–215.
- Scott WR and Davis GF (2015) *Organizations and Organizing: Rational, Natural and Open Systems Perspectives*. New York: Routledge.
- Skogstad G (2003) Legitimacy and/or policy effectiveness? Network governance and GMO regulation in the European Union. *Journal of European Public Policy* 10(3): 321–38.
- Suchman MC (1995) Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review* 20(3): 571–610.
- Torres RT and Preskill H (2001) Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation* 22(3): 387–95.
- Van der Knaap P (1995) Policy evaluation and learning: Feedback, enlightenment or argumentation? *Evaluation* 1(2): 189–216.
- Weaver C (2007) The world's bank and the bank's world. *Global Governance* 13: 493–512.
- Weiss CH (1972) Utilization of evaluation: Toward comparative study. In: Weiss CH (ed.) *Evaluating Action Programs: Readings in Social Action and Education*. Boston, MA: Allyn & Bacon, 318–26.
- Weiss CH (1998) Have we learned anything new about the use of evaluation? *American Journal of Evaluation* 19(1): 21–33.
- Weiss CH and Bucuvalas MD (1980) *Social Science Research and Decision-Making*. New York: Columbia University Press.
- Williams K, De Laat B and Stern E (2002) *The Use of Evaluation in the Commission Services*. Paris: Technopolis France.
- Wojtowicz DM and Kupiec T (2018) Reluctant to learn? The use of evaluation to improve EU cohesion policy implementation in Polish and Spanish regions. *Evidence & Policy: A Journal of Research, Debate and Practice* 14(1): 103–20.
- Zucker LG (1983) Organizations as institutions. *Research in the Sociology of Organizations* 2(1): 1–47.

Tomasz Kupiec is an Assistant Professor at the University of Warsaw and an independent evaluator. His research focuses on evaluation use and functions, and performance of evaluation systems.

Dorota Celińska-Janowicz is a human geographer specialising in urban studies and retail geography. She is an assistant professor at the University of Warsaw and a post-doc researcher at the Vrije Universiteit Brussel.

Valérie Pattyn is an Associate Professor at the Institute of Public Administration of Leiden University (The Netherlands) and is partially affiliated with KU Leuven Public Governance Institute (Belgium).