

Evaluation Policy and Organizational Evaluation Capacity Building: Application of an Ecological Framework across Cultural Contexts

Hind Al Hudib
University of Ottawa

J. Bradley Cousins
University of Ottawa

Background: Research on the role and effects of evaluation policy is limited. Some research on the policy's role in enhancing organizational evaluation capacity (EC) is beginning to accrue but to date it has been limited largely to global Western evaluation contexts.

Purpose: We employed an ecological conceptual framework arising from our own empirical research to explore the interface between evaluation policy and EC in non-western contexts. We asked—To what extent does this framework resonate across these contexts? In the selected non-Western context, what are the salient variables moderating the relationship between policy and EC in the selected contexts? Are there differences across countries?

Setting: The present research is focused on perceptions about evaluation culture and experiences in two countries situated in the Middle East and North Africa (MENA) region, namely Turkey and Jordan.

Intervention: Not applicable.

Research design: We conducted focus groups within the respective countries with a combined total of 18 participants associated with country-level voluntary organizations for professional evaluation (VOPE). Participants worked in government, non-governmental aid agencies, universities and private sector organizations.

Data collection and analysis: We introduced the focus group participants to our ecological framework and then guided the conversation using semi-structured questions. Data were audio-recorded, transcribed and subsequently thematically analyzed using NVivo.

Findings: The ecological framework was found to resonate well but the findings were weighted heavily toward macro-level contextual variables. Even though important contextual and cultural differences between Turkey and Jordan were evident, leadership emerged as a significant meso-level moderating variable in both settings. The discussion of the results included implications for ongoing research.

Keywords: *evaluation capacity building; evaluation policy; program evaluation; cultural context*

Introduction

Evaluation policy is a key component of evaluation practice that some scholars and practitioners would argue profoundly affects the day-to-day work of all evaluators and ultimately the quality of the programs they evaluate (Trochim, 2009). Trochim defines evaluation policy as “any rule or principle that a group or organization uses to guide its decisions and actions when doing evaluation” (p. 16). In essence, every organization that engages in evaluation, including governmental and para-governmental agencies, universities, not-for-profit organizations, non-governmental organizations (NGOs), and even private businesses have evaluation policies, whether explicit and written or implicit and ad hoc principles or norms (Dillman & Christie, 2017; Trochim, 2009).

Few would contest the proposition that the world is changing at a rapid and perhaps even unpredicted pace, as is evidenced by far-reaching government reforms, shifting economic conditions, financial crises, significant variance in energy prices, climate change, the degradation and depletion of natural resources and, most recently, a global pandemic. Such change has led many governmental, para-governmental organizations and NGOs to develop explicit evaluation policies and strategies with the intention of building their evaluation capacity, managing their resources, and evaluating the impact of their work (Mayne, 2007, 2009; Rist, 2006). Explicit, written evaluation policies potentially govern the ways in which organizations conduct and use evaluation and help them to realize their intention to be accountable to stakeholders, to support organizational learning and to use evidence as a basis for decision making.

Despite the recent increase in interest in evaluation policy and its impact on most aspects of evaluation practice and theory (Al Hudib & Cousins, 2020; Christie & Lemire, 2019; Dillman & Christie, 2017; Trochim, 2009), a survey of the evaluation literature shows that evaluation policy is not yet an integral part of the discourse on how to improve evaluation practice or to enhance evaluation capacity. Our targeted searches found very few conceptual papers and

empirical studies are almost non-existent (see also Christie & Lemire, 2019). Our own research program (Al Hudib & Cousins, 2019, 2020) and those of others recently working in the area (e.g., Christie & Lemire, 2019; Fierro, 2019; Kinarsky & Christie, 2019) are just beginning to explore the interface between evaluation policy and organizational evaluation capacity (EC).

Much of this research has focused on organizations located in North American and European contexts, and to a limited extent, on bilateral and multilateral organizations with a primary interest in development cooperation. Understanding the interface between evaluation policy and organizational EC in international development contexts is becoming increasingly important. This is particularly the case given the United Nations’ establishment of sustainable development goals (SDGs) in 2015 as well as the current trends in development cooperation—such as increased expectations, participation and responsibilities for recipient countries (Alonso & Glennie, 2016)—and their implications for evaluation’s role. Yet most of the scholarship in this stream of inquiry has been grounded in Western contexts, systems and thinking. Further, evaluation in the global South is born out of the needs of funding agencies (Carden, 2013; Hay, 2010) and “evaluation as conceptualized and practiced today remains very much a Western practice” (Chouinard & Hopson, 2016, p. 250).

In this paper, we want to extend our own research on the interface between evaluation policy and organizational EC to non-Western cultural contexts. Specifically, we want to explore the extent to which a three-level ecological framework linking evaluation policy and EC is applicable in two Middle-Eastern country contexts. Our research questions are as follows:

- To what extent does the framework resonate across these contexts?
- In the selected non-Western contexts, what are the salient variables moderating the relationship between policy and EC? Are there differences across these countries?

The framework (see Figure 1 and detailed description below), which emerged from our

research in the global West (Al Hudib & Cousins, 2019), illuminates the multiple interconnected variables that moderate the relationship between evaluation policy and organizational capacity to do and use evaluation. In this study, guided by the ecological framework, we explore contextual facilitators and barriers that influence the role of evaluation policy in enhancing the integration of evaluation into organizations' decision making and learning systems in two specific non-Western cultural contexts.

"There are no culture-free evaluations, as culture itself remains a socially, politically, and historically vibrant construct that is fundamentally expressive of social values and norms" (Chouinard, 2016, p. 239). We begin this article by reflecting on the meanings of cultural context and the importance of thinking about evaluation theory and practice from a cultural perspective. We then describe our ecological framework and our methods for this exploratory research study before turning to our findings. We close by discussing our findings and suggesting issues for further study.

Culture and Evaluation

Years ago, Patton (1985) compiled an edited volume on culture and evaluation in response to the increasing interest in unraveling the cultural dimensions of international evaluation practice. This concern was said to stem from the power of culture to make us relatively oblivious to the limitations of our own perspectives, behaviors and values. Patton asked "What happens when we export the ideas, concepts, models, methods, and values of evaluation to other countries and cultures?" (p. 2).

Following the publication of that volume, a great deal of research and scholarship on culture in evaluation has been published, much of it motivated by similar concerns. Even though, culture has been defined in a range of ways (both inside and outside of the evaluation field), most writers are in accordance with the definition put forward by Frierson, Hood, Hughes and Thomas (2010): "a cumulative body of learned and shared behavior, values, customs and beliefs common to a particular group of society." (p. 75). As is

noted by some scholars (e.g., Acree & Chouinard, 2019, Chouinard & Cousins, 2009) culture is a dynamic concept whose meaning has shifted over time and across disciplines; and it informs and shapes the values, philosophies and methods used by researchers and evaluators.

Much research and theory has been focused on so-called culturally responsive evaluation (CRE) which is predominantly concerned with evaluators working in cross-cultural contexts. CRE is decidedly normative in nature and implores evaluators to genuinely engage with intended program beneficiaries in the evaluation process and to assure inclusion of their values and perspectives (Chouinard & Cram, 2019; Hood, Hopson, & Frierson, 2005; Hood, Hopson, & Kirkhart, 2015). Perhaps not surprisingly, strong connections between CRE and collaborative approaches to evaluation have been made (Acree & Chouinard, 2019; Al Hudib, Cousins, Oza, Lakshminarayana, & Bhat, 2016; Askew, Beverly & Jay, 2012; Chouinard & Cousins, 2009). Nevertheless, our main concern in this paper is descriptive – that is, to explore the cultural relevance of an ecological framework in non-Western contexts and to learn what it can tell us about evaluation in these settings.

The recent review and integration of empirical evaluation work in international development conducted by Chouinard and Hopson (2016) is revealing. The authors observed that over the past 50 years, scholars and practitioners working in international development have made very limited progress in terms of integrating cultural considerations into their work. Moreover, while there has been significant progress in terms of the recognition of cultural context in evaluations conducted in North America and Europe, this recognition appears to be lacking at the global level. Development evaluation is heavily shaped by the needs, understandings, and values of international bilateral and multilateral aid agencies. Chouinard and Hopson conclude that much work is needed to ensure that non-Western perspectives, worldviews and cultures become part of the evaluation conversation.

In the present study, we start on this trail by exploring the extent to which our evidence-based ecological framework can help us understand the interface between evaluation

policy and EC in specific non-Western contexts. Is it fully, partially or not at all applicable? Through such exploration, we hope to contribute to the collective understanding of how cultural context can help shape evaluation theory and practice.

Framing the Evaluation Policy-Evaluation Capacity Interface

The two focus groups, described in detail in the methods section to follow, were framed by our ecological framework appearing in Figure 1. Though directly emerging from our prior empirical study involving 18 evaluation scholars and practitioners from Canada, the United States, and Europe with expertise in

ECB and/or evaluation policy (Al Hudib & Cousins, 2019), the development of the framework was also informed by Bronfenbrenner's ecological system theory (1979). Given its systematic approach to explaining social development and its emphasis on the importance of each level for the development process, Bronfenbrenner's theory is well suited to understanding evaluation capacity development. The theory recognizes the influence of context, community, and culture on the process of development and seeks to explain the interrelationships among multiple variables affecting that process, a perspective that helps to promote a holistic understanding.

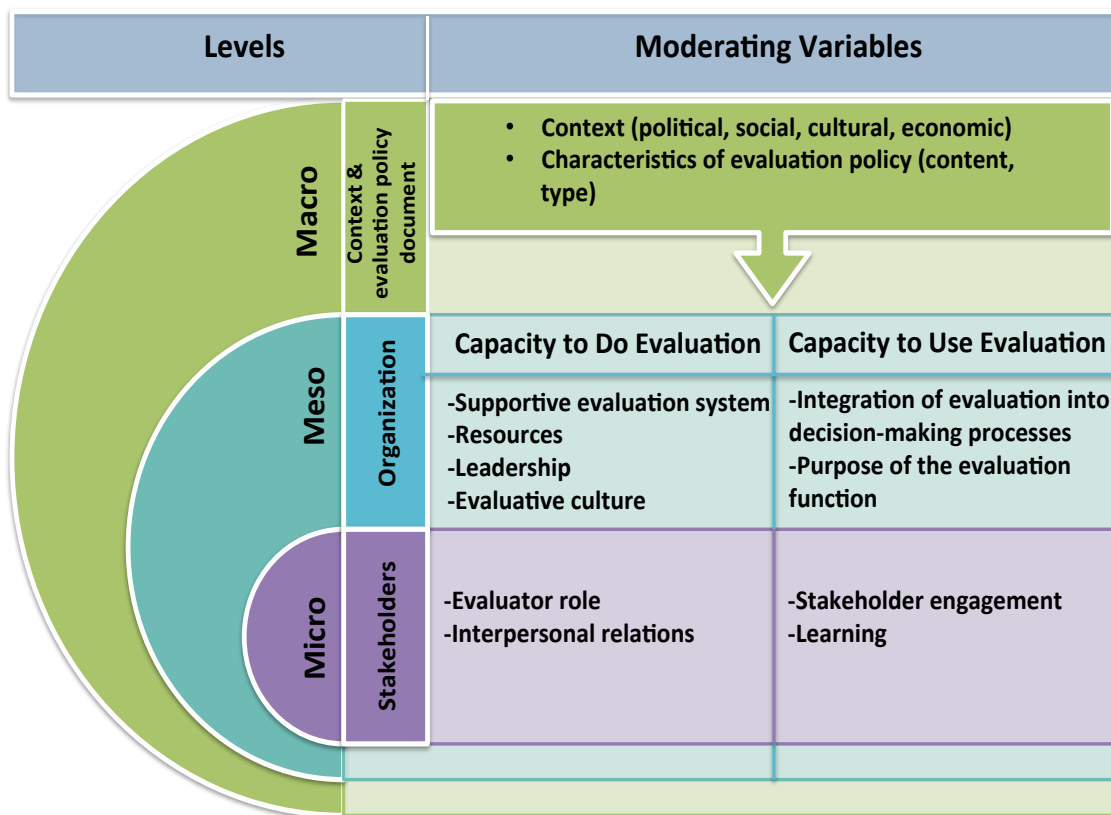


Figure 1. Ecological framework of moderating variables that explain the relationship between evaluation policy and organizational capacity for evaluation (from Al Hudib & Cousins, 2019).

In Figure 1, we can see that the framework comprises variables and key considerations at three distinct levels. First, at the *macro*-level, the focus is on the social and political context of the organization under consideration and characteristics of the evaluation policy itself. Second, the *meso*-level identifies essential moderating variables associated with organizational capacity to do evaluation (e.g., leadership, evaluative culture) and its capacity to use it (e.g., integration of evaluation into organizational decision making). Finally, at the *micro*-level evaluator and stakeholder roles and key considerations around interpersonal relations and engagement are recognized for their particular importance.

Methods

Design and Site Selection

In this study, data were collected in two semi-structured focus groups, each comprised of professional evaluators and evaluation community members from countries in the Middle East and North Africa (MENA) region. The respective focus groups took place in December 2017 in Turkey and Jordan with a combined total of 18 participants. Parallel questioning across groups permitted country-level comparisons.

The MENA region was selected for this study for a number of reasons. First, we already have some knowledge of evaluation policy and its interface with EC in the global West (e.g., Al Hudib & Cousins, 2019; Fierro, 2019; Kinarsky & Christie, 2019) where evaluation has a relatively long history. However, we currently know little about evaluation policy and its connections, if any, to EC in the MENA region. Second, by focusing on MENA, a region where evaluation capacity and practice are relatively underdeveloped, it is likely that we will gain new insights into the links between evaluation policy and EC. One of the countries (Jordan) is a developing country, while the other (Turkey) is an emerging or transitional economy, a contrast that provides interesting country-level cultural variation. Finally, there is growing interest in evaluation in the region, as is reflected in the establishment of the MENA Evaluators Network (EvalMENA) in 2011 and in a recent

conference held in Istanbul in February 2015 as part of the International Year of Evaluation global agenda. In addition, the selected countries recently established voluntary organizations for professional evaluation (VOPEs), specifically the Jordanian Development Evaluation Association (EvalJordan) in 2014 and the Turkish Monitoring and Evaluation Society (TMES) in 2013.

Study Participants

We arranged to conduct the respective focus groups with the evaluation communities in their home countries of Turkey and Jordan. Further information about each of the focus groups follows.

Turkish Focus Group: TMES. TMES was established in 2013 and was reenergized in October 2017. This provided an excellent opportunity to discuss TMES members' perspectives on evaluation policy and the ways in which these perspectives could influence their practices during the early stages of the society's development. The discussions acknowledged and were influenced by the fact that TMES aims to foster an evaluation culture in Turkey by strengthening both the supply side and the demand side of evaluation and by undertaking a series of capacity-building activities within the country. The focus group was held in Ankara, Turkey with eight participants who play an active role in Turkey's evaluation community. The participants were from various regions of Turkey but most were located in Ankara, the nation's capital. While the majority of the participants were working within government organizations, some were university professors.

Jordanian Focus Group: EvalJordan. EvalJordan was established in 2014 with the goal of improving evaluation practice in Jordan and supporting the government and civil society organizations in making informed decisions about policies and programs. We selected EvalJordan to participate in this study because of its active role in promoting evaluation practice, not only in Jordan but also in the MENA region as a whole, and

because of its continuous investment in capacity-building activities. The professional association is supported and legally recognized by the Jordanian government as a vehicle for cultivating an evaluation culture in the country and for creating a platform for evaluators and institutions that are in need of high-quality evaluations.

Jordan is a developing country in which bilateral and multilateral aid organizations are active, and the country's capital, Amman, is the location of the regional office of a number of aid and development agencies (e.g., the UNICEF regional office). The focus group session took place in Amman with 10 participants. All were members of EvalJordan and were employed by government, aid, and private sector organizations.

Data Collection and Analysis

The focus groups sessions were conducted in English, guided by open-ended questions and lasted for about 90 minutes each. At the outset, the participants were introduced to Figure 1 by way of a presentation of its origins and meaning. We then guided the conversations with questions focused on issues related to the MENA context and with questions designed to help the participants reflect on the role played by evaluation policy in fostering evaluation capacity (e.g., Does evaluation policy play a role in fostering organizational capacity for evaluation? To do evaluation? Or to use it?) and on the contextual factors that influence this role (e.g., What conditions or variables need to be in place for evaluation policy to foster capacity?). We used qualitative content analysis to identify themes and trends in the focus group data. The lead author summarized the audio recordings of the two focus groups into individual Word files that were then integrated into NVivo software as sources and codes were applied directly to them. The analysis comprised two stages. During the first stage, the lead author listened to each focus group recording and took notes on impressions, hunches, and possible themes looking for convergence across the two focus groups (Patton, 2002). Identifying a significant number of emergent themes and then exploring relationships between and within

themes. The second stage involved the classification of the themes according to their level and focused on the relationship between evaluation policy and organizational capacity for evaluation. Following Miles and Huberman (1994), we used our ecological framework as a guide. Both authors collaborated on this process in order to “support, strengthen, modify, or disconfirm the findings” (Saldana, 2009, p. 229). This part of the processes included some modest reshaping of the themes. We now present findings from our analyses associated with these key variables.

Findings

The findings confirm the relevance of the ecological framework (Figure 1) across both focus groups and adequately captures the variables that moderate the relationship between evaluation policy and EC within both countries. While the findings touch on all of the variables within the ecological framework, applicability was generally weighted to the macro-level in relation to social, political, historical and economic contextual considerations. Nevertheless, we did observe some relevance of “capacity to do evaluation” at the meso-level (organizations) and at the micro-level (stakeholders) in both countries. In addition, we observed some important differences between the Turkish and the Jordanian contexts. Notwithstanding these differences participants in both focus groups were emphatic in their confirmation of the need for evaluation policy to help build evaluation capacity and to bring people onto the same page with regard to valuing and practicing evaluation. They expressed the belief that well-thought-out evaluation policy would provide clarity around evaluation concepts, processes, principles, requirements, roles and responsibilities, use and accountability.

In what follows, we present the findings obtained from the focus groups as they relate to each level of the framework and discuss insights drawn from both focus groups in order to explain the role and use of evaluation policy as well as its influence on evaluation practice and EC. As an advance organizer, we provide a summary table associated with each

level of the framework: macro, meso, and micro.

Macro-Level: Social, Political, Historical, and Economic Context

Details about contextual aspects of Turkey and Jordan appear in the first panel of Table 1. The countries are quite distinct with Turkey having become a parliamentary representative

democracy with a highly centralized political system in 2018 (a transition that was in progress during our visit). Jordan, on the other hand, has been a centralized constitutional monarchy for many years. While Turkey is seen as an emerging economy, Jordan is acknowledged as developing country.

Table 1
Comparative Focus Group Summaries at the Macro-Level

Macro-Level (Context)	Turkish Focus Group	Jordanian Focus Group
Political, social, cultural, historical, economic contexts	<ul style="list-style-type: none"> -Significant emerging economy -Turkey is a founding member of the OECD (1961) and the G-20 major economies (1999) -Classified by the World Bank as upper-middle income country -Parliamentary representative democracy until 2018 when a new presidential system was adopted. Now, the president of Turkey is both the head of state and the head of government. -Highly centralized political system 	<ul style="list-style-type: none"> -Developing country with a relatively small economy -Jordan receives aid from many bilateral and multilateral aid agencies (e.g., USAID, World Bank, UN) -Highly centralized constitutional monarchy government
Characteristics of evaluation policy (content, type)	<ul style="list-style-type: none"> -Formal evaluation policy does not exist at the national level - Long history of audit (internal and external) in the government institutions -In 2005, the government started to focus on Results Based Management (RBM). -Turkish Court of Accounts carries out performance audits -Minimal interest in evaluation at the government level - Not dependent on international aid therefore not exposed to external evaluation demands 	<ul style="list-style-type: none"> -Formal evaluation policy does not exist at the national level -In 2006, the government initiated the <i>King Abdullah II Award for Excellence</i> at the national level to promote performance improvement and awareness and to recognize quality achievements nationwide -Strong presence of and relationship with of bilateral and multilateral aid agencies with regional offices in Amman -Dependent on international aid and strongly exposed to external evaluation demands

We learned that, in either country, evaluation policy does not exist at the national level and is not a component of any strategic planning within government institutions. Nevertheless, some Jordanian participants shared that there are monitoring and evaluation units in some government institutions and that some evaluation activities are taking place in various organizations. On the other hand, in Turkey there was a stronger emphasis on audit.

Turkish participants shared that some forms of evaluation-related activities started in the country in 2005 when the government began to move towards results-based management (RBM) and when organizations started to establish and integrate independent internal-audit systems. The Turkish participants commented that the development of fiscal and monetary policies and regulations tailored to the capital market as well as other changes in the latter were combined with more of a focus on audit practice. As a result, government organizations are required to perform independent internal audits in accordance with audit standards in order to assess compliance with current regulations and in order to demonstrate effective use of resources.

At the Turkish government level, the participants pointed out that the Turkish Court of Accounts carries out performance audits using a set of predetermined indicators in order to assess whether public resources have been used effectively. According to the participants, regardless of the designation of this type of audit as performance audit, the practice of the Turkish Court of Accounts is similar to evaluation practice. However, because the government is not particularly interested in evaluation, more restrictions have been imposed on this function. According to one of the participants,

The Court of Accounts does performance auditing. It was like an evaluation study. They ask about impact and about the effectiveness and efficiency just like an evaluation. But then the government said that you are asking these questions but

this is not the responsibility of the Court of Accounts. You are only supposed to do auditing and not to evaluate impact. So, what the government did is ... they changed the name of the function from 'performance audit' and called it 'performance information audit' so that the Court of Accounts just provides information about the indicators and prepares activity reports.

It is worth mentioning that Wilkins and Boyle's (2011) review of the standards of performance auditing used by institutions in a number of countries suggests that there remains considerable controversy about whether performance auditing is embedded in the traditions of financial auditing or is based more on evaluation traditions.

In Jordan, a driver for evaluation may be a national government initiative at the national level named the *King Abdullah II Award for Excellence*. Launched in 2006 following the establishment of King Abdullah II Center for Excellence, this award is aimed at promoting performance improvement and awareness and recognizing quality achievements nationwide. The award is the highest level of recognition for quality in Jordan and it is conferred every two years. The award's stated mission is:

To become the national reference for quality and excellence among public, private, business associations, educational service providers and non-governmental institutions, measuring their ability and efficacy in serving their individual stakeholders and their contribution in raising the competitiveness of the Jordanian economy.¹

Many participants commented that, despite the fact that the award acts as an incentive for organizations to improve the quality of their performance, and implies a role for evaluation, the process through which the award is conferred does not particularly inform practice. The award does not create a lasting commitment to continuous improvement, a commitment that would hold great potential to drive evaluation. As one of

¹ King Abdulla II Center for Excellence <https://kingabdullah.jo/en/initiatives/king-abdullah-ii-centre-excellence>

the participants mentioned, for some organizations “the end goal is the award of excellence” (our emphasis). Another participant also explained that

For the award, you have to do this and that and if you don't do it, you are not excellent or not listed. They did not build the capacity within the organization and the staff themselves.

Consequently, rather than being designed to develop sound practices and to institutionalize them within organizations, the award merely rewards the organizations and individuals that are selected.

In Turkey, the lack of interest in evaluation and the absence of demand for evaluative information for decision making at the government level form a barrier around the development of evaluation culture and capacity building, according to some participants. Evaluation is seen by senior decision and policy makers as having no use in the Turkish government system. As one of the participants commented,

The government has no intention to set up an evaluation policy. They have no desire to be accountable and they have no desire to learn from such policy because they already know everything. On the other side, there is a history of auditing in the country, either internal auditing or auditing with the Court of Accounts. That is one of the most respected institutions in Turkey. They do auditing of all of the financial institutions and departments. So there is no need for evaluation.

Some participants attributed this approach to the prevailing social values and norms, and ‘how things are done’ in the Turkish culture. One of the participants explained that the lack of interest in evaluation as a learning and improvement mechanism stems from some common cultural norms and perspectives that are based on trust – in particular, trust that the individuals who are in positions of power know their job and do not need evaluation to help them. For example,

In Turkey and in the Eastern culture, people are not interested in learning [from evaluation]. They know everything. They

have to know everything because they are the top management. If you are in this position, it means you know everything.

Participants also commented on the economic condition of Turkey and affirmed it as a critical moderating variable that influences demand for evaluation and drives the development of evaluation capacity at different levels. For example, countries that receive support from development and aid organizations are more exposed to evaluation activities that are conducted regularly to demonstrate impact and meet the donor information requirements than countries that do not receive this kind of support. As one participant put it,

Some African countries have established their evaluation system because they receive donor money but Turkey is an emerging economy not receiving any funds. So nobody is asking for independent evaluation and there is no demand for it...when there is no demand it hinders the development of evaluation.

Although requirements stipulated by donor organizations might be seen as a demand-generating factor, it should be noted that donor requirements are not a substitute for local demand for and commitment to evaluation for capacity building.

One of the Turkish participants shared an important insight about the impact of macro-level moderating variables, namely the policymaking culture, on the potential for evaluation policy establishment:

My point of view is that the policies of the country impact the evaluation. I mean the policymaking culture of the country itself. It impacts how to do evaluation, how to use evaluation and how the leadership treats the evaluation. From my international development experience, the policymaking culture impacts how we formulate policies, how we implement policies, how we evaluate policies and how this cycle is determined. Is it top down? Is it bottom up? This is very important for evaluation and the use of evaluation.

In contrast, an important macro-level consideration at the national level in Jordan is the fact that there is a strong presence of

multilateral aid agencies (e.g., UNDP, UNESCO, USAID) that are jointly working on development programs with government organizations and that have their regional office in Amman, the country's capital. The participants shared the observation that the relationship with international development agencies and their involvement in development cooperation projects have indeed influenced evaluation activities at the organizational level. As one of the participants mentioned,

I work in the Ministry of Land and International Population. It is one of the first government institutions to introduce the concept of evaluation to the government. This came from our work and relationships with the donors and international community. Learning about the evaluation function came out of our relationship with them.

Yet, the discussions with the participants revealed that there is no substantial demand for evaluative information for strategic planning or decision-making processes at the government level, though there was a definite sense that there should be. For example,

I think the government should focus on evaluation. We have a problem of corruption in Jordan. We have corrupted institutions and if we don't have evaluation, we will continue falling down.

In sum, as is the case with the Turkish focus group, the discussion with the Jordanian focus group underscores the critical influence of the macro-level variables on evaluation and ECB activities. Thus, in order to understand the relationship between evaluation policy and organizational capacity for evaluation, it is important to capture the continuous interplay among macro-level variables at the organization's external frontier that condition the role and performance of evaluation within the organization.

These general conditions at the macro-level have proven to have impacts on evaluation capacity at the meso-level.

Meso-Level: Organizational Context

Capacity to do Evaluation. As shown in Table 2, Participants in the Turkish focus group contextualized the influence of the macro-level moderating variables on organizational capacity for evaluation, especially on the capacity to do evaluation. They shared the view that leadership plays a critical role in promoting evaluation activities in Turkish organizations. For example, a number of the participants commented that, although evaluation is not mandatory in Turkey, managers could still ask for evaluation as a management practice or to inform decision-making. On the other hand, it is common for managers to resist the mandatory internal audit, not to mention the evaluation. According to one participant,

Something is missing in the culture. Some of the top management don't even want audit so they will be, like, "you are evaluating me? My performance? And you're giving this information to the Court of Accounts! You are working in my institution." This is the kind of culture. So they would come up with their own strategies to prevent the auditors from making the audit or they don't recruit them or if they have them they give them something else to do. We have this hierarchy and the management is the problem.

Table 2
Comparative Focus Group Summaries at the Meso-Level

Micro-Level (Stakeholders)	Turkish Focus Group	Jordanian Focus Group
Capacity to Do Evaluation		
-Evaluator role -Interpersonal relations	<ul style="list-style-type: none"> -Lack of demand for evaluation as leaders are perceived "to know everything" -Manager resistance to the mandatory internal audit (not to mention the evaluation) -Managers rarely ask for evaluation, even though they can -Minimal evaluative culture; evaluation mostly perceived as criticism in Turkish culture -Minimal resources to promote evaluation -Strong need for evaluation policy to build capacity 	<ul style="list-style-type: none"> -Strong emphasis on the critical role of leadership in building capacity, promoting evaluation culture and funding evaluation activities -Evaluation most often introduced in organizations by managers -Minimal evaluation culture as Arabs perceive evaluating one's work as personal criticism
<hr style="border-top: 1px dashed black;"/>		
Capacity to Use Evaluation		
-Stakeholders engagement -Learning	<ul style="list-style-type: none"> -Minimal evaluation use -Minimal interest in the "learning" function of evaluation at management levels -Common misunderstanding of the meaning of evaluation and confusion with audit 	<ul style="list-style-type: none"> -Minimal evaluation use -No decisions are based on evaluation work -Severe limitations on the sustainability of the evaluation function -Strong need for evaluation policy to build capacity

Discussion with the Jordanian focus-group members confirmed the significant role of organizational leadership in enhancing evaluation. One of the participants explained it this way:

...currently none of our projects will reach the last stage without passing through the M&E department. Yes, it exists because the leadership believes in it and there is a budget for it.

This comment illustrates the critical role of leadership in building organizational evaluation capacity; when the management is

interested in evaluation, it creates the demand, provides the needed support and makes use of evaluative information by integrating it into the decision-making process. Other participants concurred:

Leadership is more important than the other variables because they can build the other things. I mean like the budget and staff training and capacity (Participant 1).

When I look at (Figure 1), I would like to put leadership first and then leadership and then leadership. It is very important (Participant 2).

In fact, organizational leadership is a key to institutionalizing evaluation, as one of the participants indicated:

The main problem is instituting the evaluation. If we have the leadership and the policy and if we have the demand for these evaluations, we will be able to institute evaluation. We will use them.

The participants also commented on the fundamental role of cultural norms and values and how the lack of an understanding of culture influences evaluation and the perception thereof while acting as a major barrier to the development of capacity building at multiple levels. One participant explained it this way:

I think it is the culture of evaluation. We are Arab, and Arab people take things personally. People must accept and understand the value of evaluation and know that we are not evaluating them. We are evaluating the project. But people think you are not evaluating the project. You are evaluating 'my' project. There is a strong sense of ownership. Territorialism.

While this cultural element may evidently be characterised as a macro-level influence, we can see how it plays out at the organizational level.

Capacity to Use Evaluation. When asked directly about the capacity to use evaluation some Turkish participants shared the view that there is a common misunderstanding of the meaning of evaluation and that evaluation is consistently confused with auditing. The conceptualization of the meaning of evaluation has a significant impact on how people perceive its usefulness. Such perceptions are mostly connected to the reasons why evaluation was introduced in the first place. One of the participants commented that

The biggest challenge for evaluation is that evaluation is seen as equal to audit. It is used for punishing people. Making evaluation popular is difficult because of this understanding.

This long-held misconception poses a challenge for evaluation and requires a

cultural change as a prerequisite to being corrected.

The Jordanian participants' comments set out below show the level of frustration arising from the lack of evaluation use at the management level. All of the focus-group participants are of the view that an evaluation policy would inform management practice and would facilitate the integration of evaluation.

There is no appreciation of evaluators. There is no appreciation of the work they do and no decision making based on evaluation work. Nobody uses the results. We work and work and then the leadership changes or the minister changes. We evaluate for the sake of evaluation. There is a sustainability issue (Participant 1).

I think if we have a policy, the leaders will have no option. They will have to do evaluation and it will be integrated in the system (Participant 2).

This could be attributed to a general lack of awareness of the meaning and value of evaluation and of the learning benefits that it could bring to organizations. As this participant explained,

...the concept itself is an issue. No difference between evaluation and monitoring. We need to raise people's awareness about the meaning of evaluation. When we are talking about evaluation, we need to focus on the learning and impact.

The participants also believe that evaluation policy will allow for the establishment of a system that facilitates the flow of information within the organization, increases coordination among units and departments, and specifies roles and responsibilities for evaluation. According to the participants, the lack of clarity around these issues makes it very challenging for organizations to develop evaluation capacity:

We are struggling with the institutionalization in terms of how to collect data. Who will collect data and for what purposes? Therefore, we need a policy. A clear policy that has all of these components like who is responsible for

what, what do we mean by evaluation, and all of the other definitions.

This implies the participants' belief in the value of evaluation policy in terms of guiding the efforts aiming at building organization evaluation capacity.

Micro-Level: Evaluation Stakeholders

Capacity to do Evaluation. At the micro-level, many participants commented on the interpersonal relations among various stakeholders and on the ways in which they

are shaped and influenced by variables at the macro- and meso-levels, as shown in Table 3. It is evident that understanding power dynamics is an integral part of evaluation and capacity-building activities because this understanding can either enable or constrain change in the practice of evaluation. Indeed, an organization's power relations frame its conception of what evaluation actually means. Complex power dynamics may even hinder equal and genuine communication between various stakeholders.

Table 3
Comparative Focus Group Summaries at the Micro-Level

Micro-Level (Stakeholders)	Turkish Focus Group	Jordanian Focus Group
Capacity to Do Evaluation		
-Evaluator role	-Severe limits on transparency	-Strong focus on the capacity to do evaluation
-Interpersonal relations	-Lack of accountability Hierarchical organizations; authority is concentrated exclusively with the leaders	-Increasing evaluation knowledge and skills at the individual level linked to collaboration and relationships with international aid agencies
	-Top-down communication flow that does not support evaluation activities	-Local evaluators use manuals and procedures coming from aid agencies as models to build their own capacity
	-Complex power dynamics between different stakeholders	
<hr style="border-top: 1px dashed black;"/>		
Capacity to Use Evaluation		
-Stakeholders engagement	-Evaluators have an important role in promoting evaluation culture and educating stakeholders about the true value and organizational benefits of evaluation	-Limited understanding of the concept of learning and of what it means to learn from evaluation
-Learning	-Newly established TMES; some enthusiasm about evaluation; some sense of being values-driven to promote learning and bring about positive change.	-Importance of producing quality evaluation
		-Advocacy role for EvalJordan for the creation of evaluation policy.

The following discussion among three Turkish focus-group participants provides a sense of the frustration that can accompany what some stakeholders consider to be a complex web of interpersonal relationships that affect evaluation:

I think maybe in Canada there is a demand from the micro-level that drives the need for evaluation and requires accountability and transparency. In our culture, we don't have that. You can't even think of asking. I mean could you imagine asking the prime minister "How did you spend this money?" (Participant 1).

This is the job of the opposition party to ask (Participant 2).

No, it is not the job of the opposition party. It is our job. I can't even ask my municipality. It is impossible to reach them. I'm a professor and I can't reach them (Participant 1).

I think it is not the fear of asking. It is the culture that "we know everything." We also think the government knows everything, so why ask? I think it is trust or maybe because of the religion, you know, we have to obey (Participant 3).

This conversation reveals the complexities introduced by culturally grounded power dynamics and their implications for limiting transparency, accountability and communication flows.

Discussion with the Jordanian focus-group participants provided insight into the impact of working with aid agencies on the development of capacity to do evaluation at the individual level. Some shared the view that their collaboration and relationships with international aid agencies helps to foster capacity to do evaluation by increasing evaluation knowledge and skills. One of the participants shared an example to explain how working with USAID cultivated evaluation capacity at the individual level:

USAID helps with enhancing the capacity of staff working on their projects and improving the engagement in evaluation. This has a good impact on not only the project and organization level but on the national level as well because we leveraged

that here in EvalJordan and managed to build a relationship with the national evaluation department to deliver evaluation training sessions in our last conference. And we will build on that because we need the sustainability. The manuals and procedures are coming from the USAID. They will be models for us to follow as we work on building our own capacity.

The ongoing influence of international aid agencies and associated demands for accountability is undeniable.

Capacity to Use Evaluation. All participants in both countries believe that evaluators have a role in promoting evaluation culture and in educating people about the true value of evaluation and the benefits that it can bring to their organizations. The Turkish participants suggested that evaluators have to be creative and have to identify various partners with which to communicate, collaborate and initiate dialogue in order to foster a culture of evaluation. One of the participants commented that

NGOs in Turkey probably would be more experimental in developing evaluation culture than the government. I think maybe we need to look elsewhere, other than the government to develop the evaluation culture.

As mentioned above, the focus-group participants are members of the newly established TMES. They are enthusiastic about evaluation and are values-driven in terms of promoting learning and bringing about positive change. One participant put it this way:

Evaluation minus learning equals audit. Then people understand it because they don't want the learning side. We are trying to create this culture here because we want the learning side.

In Jordan, the strong focus on building the capacity to do evaluation may reflect a sense that this is more readily under the control of evaluators; for this reason it may have received more attention than the capacity to use evaluation, which depends on other variables such as leadership, as is previously

discussed. This point is evident in the following example:

We have a gap between the capacity to do evaluation and the capacity to use it. We have projects funded by USAID and they have an evaluation policy.....They keep saying use it and use the results for learning but this brings us back to the importance of leadership. They don't use it. They put it on the shelf.

This participant also commented that there is a lack of understanding of the concept of learning and of what it means to learn from evaluation, especially when working with international development agencies:

We keep hearing about the concept of learning from the donor, but we also don't know what kind of learning. They want us to learn, and they are not only focusing on the results but the problem is the evaluation capacity of the implementer.

Many of the Jordanian focus-group participants strongly expressed the view that having evaluators with adequate skills and training to design and conduct evaluations, as well as to plan for and facilitate learning from the evaluation, is essential for promoting evaluation. The following comments underscore the importance of the evaluator's role:

The problem is the evaluators. The people who are working on evaluation they don't have knowledge or clear understanding of evaluation so when they write the reports the reports are junk. They are useless and with no value. Therefore, the leaders don't take that report into consideration for decision making when they need to make a decision (Participant 1).

I think we have to decide where the problem started. Did it start with the rubbish that goes to the leaders or the leaders who don't support evaluation? (Participant 2).

We have to build the capacity and keep the pace and provide the decision makers with good evaluation so they can see the value of it and then they will use it and support it (Participant 3).

It is clear that the members of the focus group take responsibility for promoting evaluation practice and knowledge and realize that, as evaluators, they have an active role to play in fostering evaluation culture by providing credible and useful evaluation results. As members of EvalJordan, the participants also believe that their role is important for cultivating evaluation culture at the national level and so they act as advocates for the creation of an evaluation policy. Indeed,

We need to continue working on our own capacity. We need to raise awareness. We also need to identify stakeholders at the national level—people who need evaluation and will use evaluation. This will be a robust exercise. Then we need to use a change management approach and establish champions within ministries and within different levels of the government and use these champions to take this effort forward with the vision of putting in place an evaluation policy...We need a national team to develop a system.

All of the focus group participants are keen on the idea of having an evaluation policy. They feel strongly that this is what they need in order to organize their efforts and move evaluation forward at multiple levels:

We have many plans and many guidelines. Things need to be filtered. We need one unified policy so everyone can go back to it if they have any questions (Participant 1).

Yes, we need one evaluation policy. A clear policy. A clear accountability. A clear action plan so we can move forward (Participant 2).

Finally, one of the participants concluded by stating, "Let us encourage people that we are doing this for the benefit of everybody." This suggests that evaluators can play an important role in communicating the benefits of evaluation and in strengthening its usefulness and relevance to various stakeholders.

We conclude from this study that the ecological framework has merit as a lens for examining the relationship between evaluation policy and organizational EC in the MENA region. Moreover, it is entirely evident that either contextual variables are powerful influencers on this relationship either as

facilitators or, as was more to the point in the focus group data, as barriers leveraging ECB from policy. We now turn to a discussion of the findings.

Discussion

Over the course of this empirical study, the indispensable role of contextual factors and conditions and their influence on the relationship between evaluation policy and organizational capacity for evaluation became more and more evident. Our findings highlight that the socio-political, cultural and economic context of an organization is critical and has pervasive influences on evaluation, which is consistent with the recent work of scholars who call attention to the importance of context as the most salient influencing and shaping feature of evaluation practice (e.g., Chouinard & Milley, 2016; Coldwell, 2019; Vo & Christie, 2015). Our findings are also aligned with Pawson and Tilley's (1997) approach to context as the social and cultural conditions within which programs, initiatives or interventions occur. These conditions encompass both the structural setting (i.e., organizational, spatial and temporal), and the individuals involved, including their personal characteristics and interpersonal relationships, which Pawson (2013, p.37) referred to as the 4 I's: individuals, interpersonal relations, institutional settings and infrastructure (the cultural, economic and social aspects of the setting). Additionally, as Pawson (2016) recently noted:

Context is layered. Sometimes it is pre-existing, macro-economic conditions that need to be auspicious to forward a policy. Sometimes it is institutional norms that need to be supportive to enable change. Sometimes it is cultural practices that need to be consonant with a new program. Sometimes it is the prevailing interpersonal relations that need to be favorable for an intervention to work (p. 49).

From the perspectives of international development (Raimondo, 2018) and culturally responsive evaluation (Acree & Chouinard, 2019), scholars have discussed several gaps in the literature pertaining to the use and influence of evaluation. They have emphasized the fact that the existing evaluation theories in

these areas are based on models of rational organizations that ignore contextual factors like institutional norms, routines and belief systems. The Raimondo further argues that important insights from international organization theory allow for better understanding of complex conjunctions of context, including those of material, cultural, internal and external factors affecting evaluation, on the one hand and of processes of change at the organizational and environmental levels on the other hand. In their review and integration, Acree and Chouinard (2019) concluded that "the concept of use...could not capture the dynamics and complexities of culturally responsive practice in [their] sample of studies" (p. 12).

The participants in present study reflected on how governments, social policies, legislation, national economic conditions and culture are all contextual factors that moderate both the role of evaluation policy in building organizational EC and the ways in which people perceive evaluation in their respective countries. The most substantive focus was on the moderating variables related to the capacity to do evaluation at the organizational and individual levels. Contextual variables touching on experience with evaluation, perceptions of its worth and fit, and evaluation demand, were shown to be quite influential. Even though the countries are from the same region, they are remarkably distinct from one another and, as such, the macro-contextual issues played out in different ways. An interesting difference between Turkey and Jordan is Turkey's lack of evaluation demand inasmuch as it is an emerging economy and therefore is not reliant on donor or international-development funding. Jordan, on the other hand, has a considerable history of working with external donors and therefore has experience with evaluation.

Based on the findings of this study, it can be argued that, in the MENA region, the very concept of evaluation is problematic. The fundamental reason for this is simply that the idea of evaluation as it is perceived and understood in the West does not align very well with the Middle Eastern culture where preserving dignity and 'saving face' are more significant and have a higher value than evaluating and judging one's work. As was

evident in our study, Middle Eastern people in their respective societies are expected to show personal integrity in order to be socially accepted. They manifest honour and integrity by projecting a public impression of dignity derived from an ostensible lack of guilt. Even if the facts and conditions point to the contrary, the social image of non-guilt must be preserved if they are to maintain the socially expected face. Dignity and respect are granted to those who show themselves to be flawless; in general, Middle Eastern societies accord no respect to people whose faults or errors come to public knowledge. Within this culture, blame, fault or error is likely to cause a fall from social grace and a loss of dignity or face; thus, Middle Eastern people generally feel revulsion and bitterness towards anything that might tend to compromise them in this way, including evaluation (Fairholm, 1994; Greaves, 2008; Sarayrah, 2004). For this reason, within this cultural context it is critical for evaluation policies to frame evaluation in a way that emphasizes its learning benefits rather than mainly focusing on accountability.

Failure to recognize and fully understand the various contextual factors that support or impede capacity-building efforts can result in an evaluation policy that inadvertently undermines stakeholders' ownership of the development process by creating a vicious cycle of disempowerment and a vicious circle of demotivation within organizations (Theisohn & Courtnadge, 2005). For example, the use and abuse of power constitute one of the aspects of context that arguably is the most important to evaluation policy but also is the most seldom analyzed in the evaluation literature. Evaluation policies that are solely aimed at enhancing accountability and legitimacy may actually exacerbate entrenched power-differential issues, especially in organizations operating in contexts of monarchical, non-democratic or dominant-party governments. In these contexts, as was proven to be the case in Jordan and Turkey, there is a long history of hierarchy as the natural order of things; high power distance is one of the principal features of the relationship between officials and peasants, rulers and ruled, and elites and masses. Therefore, a lack of understanding of these contextual issues may result in

evaluation policies that support the preservation of the existing power imbalance, and this in turn can have a negative impact both on evaluation and on ECB processes. Thus, the emergent nature of the capacity-building process within these contexts highlights the need for evaluation policies that address complex systems and that encourage flexibility, adaptability and innovation.

It is clear that for evaluation policy to have a beneficial influence on the development of organizational EC, especially in non-Western, developing-country contexts, thorough contextual analysis should be recognized as a precondition for the provision of effective support to organizational ECB processes. Such analysis can help with developing and disseminating understanding, first, of where prevailing gaps exist in the organizational capacity to do evaluation (including in human resources, organizational and administrative systems) and, second, of why these gaps exist in relation to socio-political, economic and cultural factors, and the specific statutory and regulatory systems that constrain or enable capacity-building efforts. This means that evaluation policies that are based on the premise that capacity building is basically a matter of 'replicating best practices' regardless of organizational context are not effective and are even problematic, especially in organizations that lack the capacity to support ECB activities.

Leadership is a critical variable in cross-cultural contexts that moderates the role of evaluation policy in enhancing organizational EC and forms a link between the capacity to do evaluation and the capacity to use it at the organizational level. It is entirely evident that the various contextual factors, including the socio-political, cultural, historical and economic characteristics of a country, have a direct influence on the leadership style of organizations within the national context. Within Middle Eastern societies in general—and as was most evidently the case in our Turkish focus group—a leader is perceived to be someone who “knows it all” because of their position, expertise and achievements. A leader has all of the power to make decisions; in fact, it is likely that most people would perceive leaders as being incompetent or even weak if they overly engage with other organization members in the decision-making process.

Instead, leaders are expected to maintain a consultative approach whereby they consult other organizations' members informally but make their own decisions unilaterally in order to show their power. Such power is associated with their status and position as leaders within the organization (Jackson, 2016; Mansur, Sobral & Goldszmidt, 2017).

It should be noted that, in the Middle East and in many developing country contexts, leaders are highly status conscious. They are motivated to (and often can) remain in a position of power to maintain their status in the society. Therefore, they are most likely to resist change in order to ensure that they do not lose power or relinquish authority. It is interesting to note that this leadership style is focused predominantly on the work that leader expects and rarely on people at large. This factor, in and of itself, is more likely to hinder capacity building efforts. Given this instrumental role of leaders within organizations in these contexts, Western models and styles of leadership that are based on and are developed for democratic societies are largely inapplicable in such a highly hierarchical system (Jackson, 2016).

The presence of supportive and committed leadership to facilitate and manage the implementation of ECB efforts and to internalize norms and ownership of the process to ensure sustainability is the main element in the black box of what drives capacity building within these organizations, especially in developing-country contexts. However, it should be noted that the factors that generate this leadership themselves constitute yet another black box; the individual evaluation capacity of leaders is clearly yet another salient feature of the process.

Implications for Research

We are content that our ecological framework remained relevant across cultural contexts and we feel that this research is breaking a new ground in the field. In order to develop a clearer understanding of the role played by evaluation policy in organizational EC, we would benefit from access to a wider range and volume of studies. Additional research on evaluation policy would provide a more solid

basis for understanding variables and patterns that can provide insight into the mechanisms for capacity development, within organizations, in different contexts.

The complexity of the role played by evaluation policy in organizational EC requires continued research that goes beyond exploring evaluation community members' perspectives at the individual level and that reaches toward a better understanding of evaluation policy at an organization or system level. As the findings of this research are based on the points of view of the evaluation community members and practitioners, future studies could investigate how the ecological framework resonates with evaluation users, such as, for example, organization or program managers and higher-level decision makers. These studies could also focus on one or more particular types of organization (e.g., governmental, not-for-profit organizations) in order to generate insights related to issues concerning ECB and related to the major variables moderating the role of evaluation policy in EC in those environments.

As was made clear in this research, ECB takes place in complex environments. Thus, more research needs to be undertaken to explore the ways in which particular moderating variables, such as leadership, interact with explicit evaluation policies; how these interactions influence evaluation practice and use, especially in cross-cultural contexts, is of high interest.

Acknowledgements

This paper is based on the doctoral dissertation of the first author at the University of Ottawa. The authors are grateful for financial support provided by the Saudi Cultural Bureau. Opinions and assertions made do not necessarily reflect policy of the sponsoring agency.

References

- Acree, J., & Chouinard, J. (2019). Exploring use and influence in Culturally Responsive Approaches to Evaluation: A review of the empirical literature. *American Journal of*

- Evaluation*, 1-5. doi: 10.1177/109821401987950
- Al Hudib, H., Cousins, J. B., Oza, J., Lakshminayarana, U., & Bhat, V. D. (2016). A cross-cultural evaluation conversation in India: Benefits, challenges and lessons learned. *Canadian Journal of Program Evaluation*, 30(3), 329-343. doi: 10.3138/cjpe.30.3.06
- Al Hudib, H., & Cousins, J. B. (2019, submitted). Understanding evaluation policy and organizational capacity for evaluation: An interview study. *Under review for publication consideration*.
- Al Hudib, H., & Cousins, J. B. (2020, forthcoming). Evaluation policy and organizational evaluation capacity building In M. M. Mark & N. Hart (Eds.), *Evaluation Policy, New Directions for Evaluation*. Thousand Oaks: SAGE.
- Alonso, J., & Glennie, J. (2016). What is development cooperation? 2016 *Development Cooperation Forum Policy Briefs* (1), 1-5.
- Armytage, L. (2011). Evaluating aid: An adolescent domain of practice. *Evaluation*, 17(3), 261-276. doi: 10.1177/1356389011410518
- Askew, C., Beverly, M., & Jay, L. (2012). Aligning collaborative and culturally responsive evaluation approaches. *Evaluation and Program Planning*, 35, 552-557. doi: 10.1016/j.evalprogplan.2011.12.011
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.
- Carden, F. (2013). Evaluation, not development evaluation. *American Journal of Evaluation*, 34(4), 576-579. doi: 10.1177/1098214013495706
- Chouinard, J. (2016). Introduction: Decolonizing international development evaluation. *Canadian Journal of Program Evaluation*, 30(3), 237-247. doi: 10.3138/cjpe.30.3.01
- Chouinard, J., & Cram, F. (2019). *Situating culturally responsive approaches to evaluation: Empirical implications for theory and practice*. Thousand Oaks: SAGE.
- Chouinard, J. A., & Cousins, J. B. (2009). A review and synthesis of current research on cross-cultural evaluation. *American Journal of Evaluation*, 30(4), 457-494. doi: 10.1177/1098214009349865
- Chouinard, J., & Hopson, R. (2016). A critical exploration of International Development Evaluation. *Canadian Journal of Program Evaluation*, 30(3), 248-276. doi: 10.3138/cjpe.30.3.02
- Chouinard, J., & Milley, P. (2016). Mapping the spatial dimensions of participatory practice: A discussion of context in evaluation. *Evaluation and Program Planning*, 54, 1-10. doi: 10.1016/j.evalprogplan.2015.09.003
- Christie, C., & Lemire, S. (2019). Why evaluation theory should be used to inform evaluation policy. *American Journal of Evaluation*, 40(4), 490-508. doi: 10.1177/1098214018824045
- Coldwell, M. (2019). Reconsidering context: Six underlying features of context to improve learning from evaluation. *Evaluation*, 25(1), 99-117. doi: 10.1177/1356389018803234
- Dabelstein, N., & Patton, M. Q. (2013). Lessons learned and the contributions of the Paris Declaration: Evaluation to evaluation theory and practice. *Canadian Journal of Program Evaluation*, 27(3), 173-200.
- Dillman, L., & Christie, C. (2017). Evaluation policy in a non-profit foundation: a case study exploration of the Robert Wood Johnson foundation. *American Journal of Evaluation*, 31(1), 60-79. doi: 10.1177/1098214016642864
- Greaves, S. (2012). A primer of Middle Eastern leadership culture. *Journal of Strategic Security*, 5(4), 99-118. doi: 10.5038/1944-0472.5.4.7
- Fairholm, G.W. (1994). *Leadership and the Culture of Trust*. Praeger, Westport, CT.
- Fierro, L. (2019, November). *Research on federal evaluation policies*. Paper presented at the annual meeting of the American Evaluation Association, Minneapolis.
- Frierson, H., Hood, S., Hughes, G., & Thomas, V. (2010). A guide to conducting culturally responsive evaluations. In J. Frechting (Ed.), *The 2010 user-friendly handbook for project evaluation*. Alexandria, VA: National Science Foundation. doi: 10.1187/cbe.10

- Hay, K. (2010). Evaluation field building in South Asia: Reflections, anecdotes, and questions. *American Journal of Evaluation*, 31(2), 222-231. doi: 10.1177/1098214010366175
- Hood, S., Hopson, R., & Frierson, P. (Eds.). (2005). *The role of culture and cultural context: A mandate for inclusion, the discovery of truth and understanding in evaluative theory and practice*. Greenwich, CT: Information Age Publishing.
- Hood, S., Hopson, R., & Kirkhart, K. (2015). Culturally responsive evaluation: Theory, practice and future implications. In K. Newcomer, H. P. Hatry & J. S. Wholey (Eds.), *Handbook of Practical Program Evaluation (4th Ed.)*. (pp. 281-317). Hoboken, NJ: Wiley.
- Jackson, E. T. (2016). Paternalistic leadership: The missing link in cross-cultural leadership studies? *International Journal of Cross Cultural Management*, 16(1), 3-7. doi: 10.1177/147059581663770
- Kinarsky, A., & Christie, C. (2019, November). *Evaluation policy in philanthropies: What are they? Who has them? Why write them?* Paper presented at the annual meeting of the American Evaluation Association, Minneapolis.
- Mansur, J., Sobral, F., & Goldszmidt, R. (2017). Shades of paternalistic leadership across cultures. *Journal of World Business*, 52(5), 702-713. doi: 10.1016/j.jwb.2017.06.003
- Mayne, J. (2007). Challenges and lessons in implementing Results-Based Management. *Evaluation*, 13(1), 87-109. doi: 10.1177/1356389007073683
- Mayne, J. (2009). Results management: Can results evidence gain a foothold in the public sector? In O. Reiper, F. Leeuw, & T. Ling (Eds.), *The evidence book* (pp. 117-149). New Brunswick, NJ: Transaction.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks: Sage.
- Patton, M. Q. (1985). Editor's Notes *Culture and evaluation*, *New Directions for Program Evaluation*, No 25 (pp. 1-2). San Francisco: Jossey-Bass. doi: 10.1002/ev.1380
- Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oaks, CA: SAGE
- Pawson, R. & Tilley, N. (1997). *Realistic Evaluation*. London: SAGE.
- Pawson, R. (2013). *The Science of Evaluation: A Realist Manifesto*. London: SAGE.
- Pawson, R. (2016). The ersatz realism of critical realism: A reply to Porter. *Evaluation*, 22(1), 49-57. doi: 10.1177/1356389015605206
- Raimondo, E. (2018). The power and dysfunctions of evaluation systems in international organizations. *Evaluation*, 24(1), 26-41. doi: 10.1177/1356389017749068
- Rist, R. C. (2006). The "E" in monitoring and evaluation: Using evaluative knowledge to support a results based management system. In R. C. Rist & N. Stame (Eds.), *From studies to streams: Managing evaluative systems* (pp. 3-22). New Brunswick, NJ: Transaction.
- Sachs, J. D. (2012). From Millennium Development Goals to Sustainable Development Goals. *The Lancet*, 379 (9382), 2206-2211. doi: 10.1016/S0140-6736(12)60685-0
- Sarayrah, Y. K. (2004). Servant leadership in the Bedouin-Arab culture. *Global Virtue Ethics Review*, (5) 3, 50-58.
- Theisohn, T. & Courtnadge, P. (2005). Moving beyond the "Munchhausen Approach". Honest brokering and independent monitoring in development partnerships, in Nord-Sud aktuell, 19, 246-258.
- Trochim, W. M. K. (2009). Evaluation policy and evaluation practice. *New Directions for Evaluation*, No. 123, 13-32. doi: 10.1002/ev.303
- Vo, A. T., & Christie, C. A. (2015). (Eds.). Advancing research on evaluation through the study of context. *New Directions for Evaluation*, No. 148, 43-55. doi: 10.1002/ev.20156
- Wilkins, P., & Boyle, R. (2011). Standards and qualities. In J. Lonsdale, P. Wilkins, & T. Ling (Eds.), *Performance audit: Contributing to accountability in democratic government*, 147-173. Cheltenham, UK: Edward Elgar.