http://www.diva-portal.org

This is the published version of a paper published in *Evaluation*.

Citation for the original published paper (version of record):

Petersén, A., Olsson, J. (2026)
Universal evaluation ethics? A comparative study of ethical guidelines from an Aristotelian perspective
Evaluation, : 13563890261486831
https://doi.org/10.1177/13563890261486831

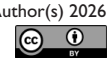
Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:oru:diva-132093

Check for updates

Article

Evaluation
1–21
© The Author(s) 2026

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/13563890261486831
journals.sagepub.com/home/evi
Sage

Universal evaluation ethics? A comparative study of ethical guidelines from an Aristotelian perspective

Anna Petersén[1] and Jan Olsson[2]

Abstract

This article examines how five international and regional ethical guidelines for evaluation articulate and balance universal ethical principles with contextual, cultural, and sociopolitical considerations. Using Aristotelian ethics as an analytical framework, this study compares the guidelines through six concepts: *telos*, *eudaimonia*, *aretê*, *phronesis*, *mesotes*, and *zoon politikon*. This analysis shows that while all guidelines share core values such as integrity, justice, and respect, they differ in how explicitly they address contextual influences and societal purposes. Some emphasize professional norms, whereas others foreground cultural relevance, community participation or global responsibility. These variations reveal underlying normative tensions in contemporary evaluation ethics and illuminate the challenges of formulating ethical principles that are both broadly applicable and contextually meaningful. This study contributes to ongoing debates on whether universal evaluation ethics is achievable and highlights the value of dialogue and contextual sensitivity in ethical guideline development.

Keywords

ethical guidelines, evaluation ethics, Aristotelian ethics, universalism, comparative analysis

Introduction

There is an ethics trend within the evaluation field in terms of increasing numbers of ethical guidelines and codes of ethics that seek to guide and support evaluators in a more professional ethical role and to integrate ethics in evaluation processes. Although not a new, dominant

[1]School of Behavioural, Social and Legal Sciences, Örebro University, Sweden
[2]School of Humanities, Education and Social Sciences, Örebro University, Sweden

Corresponding author:
Anna Petersén, School of Behavioural, Social and Legal Sciences, Örebro University, 701 82 Örebro, Sweden.
Email: anna.petersen@oru.se

evaluation wave (Vedung, 2010, 2025; Olsson and Petersén, 2025), evaluation ethics seems to be increasingly important, particularly when considering serious efforts among evaluation scholars to contribute to this process in constructive ways (Morris, 2011, 2015; Onyura et al., 2023; Schwandt, 2007, 2018; Simons, 2006; Van den Berg et al., 2022). A key issue evoked by this development is whether far-reaching ethical standardization is possible and valuable for making "good" and well-grounded evaluations in the eyes of various stakeholders, including the evaluators themselves (Onyura et al., 2023). There are some basic, reasonable arguments both for and against the ambition to strive for universal ethical guidelines.

It is in the nature of the matter that ethical guidelines are based on at least a couple of values and behavioral standards claimed to be universal, even though contextual and cultural aspects are usually acknowledged as well. A good illustration of this is IDEAS (International Development Evaluation Association), which points out that its code of ethics was developed through intercultural dialogue among members of the association, representing 90 countries all over the world, and that general use of the code is encouraged also among non-members of the association (IDEAS, 2014: 1). At least some amount of general standardization seems to be necessary for ethical guidelines to be perceived as useful and legitimate. If ethical guidelines are not institutionalized into appropriate ethical behaviors, they face the risk of functioning as nothing more than ethical "window-dressing," which would likely lead to loss of credibility within a relatively short period of time.

As a contrast, some evaluation scholars stress the importance of the practical context in making good and well-grounded judgments about values and facts since evaluations do not take place in either an abstract scientific universe of epistemic knowledge production (Mark et al., 1999) or in a world of "objective" ethical standards (Hood et al., 2015). Thus, values and evidence in evaluation processes tend to vary between different cultural, social, and political contexts (Aronsson and Hassnain, 2021; Hood et al., 2015). The same ethical principle, for example, "to promote partnership and collaboration among evaluators and related professionals with whom they work" (IDEAS CODE OF ETHICS 2013/2014, IDEAS: 3), can be interpreted and implemented differently in countries or organizations with varying cultures of participation and collaboration. The specific ethical principle can thus get different meanings in different contexts, which can make sense but also cause trouble for those who seek to develop universal ethical standards for evaluators. As Schwandt (2007) argues, many ethical guidelines in evaluation are historically rooted in Western moral assumptions, emphasizing individual autonomy, meritocratic ideals and personal responsibility. Such normative foundations cannot be taken for granted in all cultural contexts, and the meaning of acting ethically therefore shifts depending on local sociocultural norms. This reinforces the need to interrogate whether universal ethical guidelines can truly account for the diversity of evaluative environments worldwide.

This tension is addressed by both evaluation scholars (Dahler-Larsen, 2013; Schwandt, 2007) and ethical philosophers (Fleischacker, 1999; Kuçuradi, 2014; Lang, 2020). Different ideas of how to manage this tension are discussed in the literature. The "human rights" model is largely a top-down strategy, which seeks to spread and to some extent impose ethical standards and instruments, while a less ambitious strategy is the "necessary conditions model," likely delivering a "thin universal ethic" (Fleischacker, 1999: 105). A third strategy is labeled the "cultural dialogue model," which is portrayed as a serious way of expanding the conversation on evaluation ethics, which can develop cross-cultural understanding and agreements (Fleischacker, 1999; Schwandt, 2007). Several scholars have pointed out that universal ethical

guidelines risk becoming too abstract to offer meaningful direction in culturally complex or practice-based contexts. Msoroka and Amundsen (2018), for example, demonstrate how the application of ethical principles in real research settings often requires negotiation between universal norms and local cultural expectations. They therefore argue for a hybrid model, "universality with diversity," that acknowledges the limits of purely top-down ethical frameworks and allows context-specific interpretations to emerge. This perspective reinforces the need for ethical principles to be developed not only through overarching standards but also through bottom-up processes grounded in lived practice. Such an approach resonates with the core concerns of this article, namely how ethical guidelines must be both principled and contextually relevant if they are to function effectively in diverse evaluative environments.

Thus, arguments for both universal and culturally sensitive evaluation ethics make sense, and they are to some extent contradictory, likely in need of being balanced in relation to each other to achieve both ethical guidance and practical, contextual relevance. This tension is the point of departure for the analytical problem of this article; that is, how do guidelines of evaluation ethics seek to balance general ethical principles with practical, contextual relevance? The theoretical framework applied in this article is Aristotelian ethics, which is suitable since it offers a relatively broad perspective on ethics combined with practical relevance, particularly acknowledging virtue ethics as an important value as well as a source for pragmatic ethical improvement over time. Of great importance to the analytical problem is that Aristotelian ethics perceives humans as social and political animals, which means that ethics requires both individual and collective responsibilities on ethical matters (Aristotle, 2012; Pagan, 2008). Aristotelian ethics can thus work as an appropriate framework for studying ethical guidelines of evaluation ethics. Empirically, this article contributes by making a comparative study of five prominent guidelines of evaluation ethics from different parts of the world. This design allows for social, cultural, and political aspects in different contexts to influence the content of the guidelines, and thus to likely limit the possibility for universal evaluation ethics, despite general ethical ambitions. While Aristotelian ethics is not the only possible framework for examining evaluation ethics, it provides a useful analytical lens for the present study. Contemporary discussions of evaluation ethics often revolve around tensions between universal principles and contextual judgment, between procedural standards and ethical practice, and between individual responsibility and collective societal obligations (Schwandt, 2007, 2018; Van den Berg et al., 2022). Aristotle's conceptual vocabulary is especially suited to illuminate these tensions because it connects moral character, practical judgment, social and political responsibility, and broader purposes of human flourishing. The framework is therefore not employed as a normative benchmark against which the guidelines are judged as better or worse. Rather, it functions as an analytical lens that renders visible different ways in which ethical guidelines conceptualize ethical action, societal purpose, contextual adaptation, and professional responsibility.

The purpose of this article is to examine how a diverse set of international and regional ethical guidelines for evaluation articulate and balance universal ethical principles with contextual, cultural, and sociopolitical considerations. Using Aristotelian ethics as an analytical framework, this study seeks to identify shared values, points of divergence, and the normative tensions that arise when ethical ideals are operationalized in practice. In doing so, this article aims to contribute to a deeper understanding of whether, and to what extent, universal evaluation ethics is possible.

Theory

Aristotle (384–322 BC) is well-known for questions regarding the ethics, politics, and life goals of human beings. Here, we use a selection of Aristotle's concepts as analytical tools for new perspectives on ethical guidelines and how they can be applied in practice. The choice of Aristotle is motivated not primarily by historical influence but by analytical suitability. Evaluation ethics concerns how evaluators should act in situations characterized by competing values, contextual variation, and uncertainty. Alternative ethical frameworks could also have been employed. A deontological perspective would primarily direct attention toward rules, duties, and compliance, whereas a consequentialist perspective would focus on the outcomes generated by evaluation practice. Both perspectives provide important insights and have strongly influenced public administration and policy analysis. However, the central concern of this article is how ethical guidelines balance universal principles with context-sensitive judgment and collective responsibility. Aristotelian virtue ethics is particularly useful because it integrates questions of purpose, character, practical judgment, balance, and social responsibility within a single conceptual framework. Rather than focusing exclusively on rules or consequences, it enables analysis of how evaluators are expected to exercise ethical judgment in complex social settings. This integrated perspective is especially relevant when examining whether evaluation ethics can be both universal and context-sensitive.

The first concept is *telos*. According to Aristotle, telos is an aim or an inner direction. It refers to an end, fulfillment, completion, or goal. The telos of humans is the starting point for understanding how they pursue the "good life" or *eudaimonia* (human flourishing). It represents the purpose for which something is done or occurs. This means that human actions are understood as being directed toward a specific telos, and practical deliberation involves identifying the concrete steps needed to achieve it (MacIntyre, 2007 [1981]). The second concept, which has already been mentioned, is eudaimonia. It refers to the state of having an objectively desirable life, which differs from the modern term "happiness," which refers to having a subjectively satisfactory life. It comes to expression as "living well" or "human flourishing." According to Aristotle, eudaimonia cannot be reached only by developing one's virtues but also by being accompanied by external goods. Further, eudaimonia is not a temporary state of mind, but a way of living where every part of life, including actions, habits, values, and decisions, is related. One can reach eudaimonia by living according to virtue and developing one's human abilities. *Aretê* is the third concept in the analytical framework and is here translated as "virtue." Aristotle distinguishes between different virtues, but in common for all of them is that they represent excellence in human abilities. A person with excellent human abilities consequently acts and strives toward "the good," instead of just occasionally acting correctly. Somewhat simplified, aretê denotes being a good person; however, philosophers have questioned whether it is a condition of intellect or character and if one can be taught goodness or if it is inherited. To live a life according to virtues, the fourth concept, *phronesis*, is the way. Phronesis resembles prudence and practical wisdom on how to act in specific situations, general truths notwithstanding (Field and MacIntyre Latta, 2001; Tsang, 2007). Some definitions connect the concept to moral sensitivity, emphasizing that fostering "the good in life" requires a commitment to making ethically sound choices. The capacity for such decision-making is closely linked to the phronetic virtues. Practical wisdom is grounded in an understanding of what is best for the specific person in the unique situation. It entails both the ability to discern how to achieve a particular goal and the capacity to reflect upon and determine which goals

are genuinely good. For instance, phronesis becomes concrete when an experienced person understands when to disregard general rules and instead uses experience and value-based judgment to discern how to act. Thus, phronesis is primarily relevant under concrete conditions and stresses sensitivity to the unique situation (Petersén and Olsson, 2015). Schwandt's (2007) understanding of cultural competence, as a practice of reflexive engagement with the values, assumptions, and social structures embedded in evaluative work, aligns closely with Aristotelian notions of phronesis. Both perspectives reject rule-following in favor of situated ethical judgment grounded in dialogue and openness to alternative interpretations. For evaluators, phronesis implies more than compliance with ethical rules. It refers to the capacity to exercise sound judgment when facing competing values, contextual constraints, and conflicting stakeholder interests. In this sense, practical wisdom is not opposed to ethical principles but concerns their thoughtful interpretation and application in particular circumstances. The fifth concept of relevance here is *mesotes*, the doctrine of the mean. According to Aristotle, virtues are about finding mesotes, a point of balance between two extremes (Evrigenis, 1999). An example is the virtue of courage located in between cowardice and recklessness. However, in between does not refer to the middle, because it must be regarded in relation to the unique situation, which once more highlights the importance of phronesis. In the *Nicomachean Ethics*, Aristotle distinguishes between the mean in relation to one another and the mean in relation to us and argues that the latter is the best, according to science and prudence. Therefore, one ought to strive for the mean of us in our efforts to find out how to act. The sixth and final concept is *zoon politikon.* According to Aristotle, human beings are by nature political animals, that is, social beings whose flourishing depends on living within and contributing to a community (*polis*) (MacIntyre, 2007 [1981]). This idea emphasizes that ethical life is not an isolated pursuit but arises within social relations and shared practices. To live well, one must therefore act in ways that sustain and enhance the common good. The concept of zoon politikon also implies that moral development and the exercise of virtues occur through participation in collective life: through dialogue, cooperation, and engagement in public affairs. Within the context of evaluation practice, this perspective highlights that ethical judgment cannot be reduced to individual reasoning alone. Rather, it involves responsiveness to the values, norms, and purposes of the communities in which evaluations are conducted. As such, zoon politikon complements phronesis and mesotes by situating moral agency within the social and political conditions that make ethical deliberation both possible and meaningful.

Aristotle's ethical and political philosophy can be understood as an integrated system in which the concepts build upon one another. The starting point is that everything has a telos, a purpose. For human beings, this ultimately means striving toward eudaimonia—living well and realizing one's potential as a human being. To achieve this, one must cultivate aretê (virtue), the capacity to act in a way that embodies excellence. Virtue, however, is not attained through rules or abstract theory alone, but through phronesis, the practical wisdom that enables one to discern what is right in each concrete situation. The principle of mesotes also plays a central role: Virtue often lies in the balance between two extremes, and it requires practical judgment to find this balance. Finally, Aristotle shows that humans cannot realize their telos in isolation. As a zoon politikon, a fundamentally social being, one can only develop virtue and achieve eudaimonia within a community, in political and social life. Together, these concepts form a coherent whole, uniting human purpose, the path to achieve it, and the social conditions necessary for attaining the good life in an integrated theory. For the purposes of this study, the framework can also be understood in more practical terms. The six concepts capture

different dimensions of evaluation ethics: purpose, well-being, professional virtues, context-sensitive judgment, balance, and collective responsibility. Together, they provide a coherent basis for analyzing how ethical guidelines define good evaluation practice. While more straightforward categories such as goals, values, competencies, and societal responsibilities could have been used, the Aristotelian concepts add analytical depth by connecting these dimensions within a broader account of human flourishing and moral character. In particular, eudaimonia and aretê make it possible to examine not only what evaluators are expected to do, but also what forms of professional excellence and societal well-being the guidelines implicitly promote.

Method

This study employs a qualitative, interpretive research design to examine how ethical guidelines in evaluation articulate underlying values, operational principles, and cultural assumptions. Rather than aiming for a comprehensive global mapping, this study conducts an in-depth comparative analysis of five influential guidelines selected strategically to capture variation in ethical traditions and sociopolitical contexts. The empirical material consists of ethical frameworks from the American Evaluation Association (AEA), Australasian Evaluation Society (AES), Canadian Evaluation Society (CES), African Evaluation Association (AfrEA), and the United Nations Evaluation Group (UNEG). These organizations were chosen because they are widely recognized within the evaluation community and represent diverse institutional and cultural settings in which evaluation ethics are formulated and practiced.

Analytical framework and procedure

The analysis is grounded in Aristotelian ethics, operationalized through a coding scheme developed specifically for this study. The scheme translates six central Aristotelian concepts (*telos*, *eudaimonia*, *aretê*, *phronesis*, *mesotes*, and *zoon politikon***)** into analytical questions and textual indicators. These concepts were chosen because they allow an examination of ethics not only as rule-following but as value-laden, context-dependent, and socially embedded practice.

The coding scheme (Table 1) specifies the following:

- Analytical questions (e.g. “What goal is formulated for the guidelines?”);
- Operational indicators (e.g. references to justice, flourishing, balance, community);
- Conceptual definitions drawn from Aristotelian ethics.

This framework enables a systematic yet interpretive comparison across documents. The analytical process unfolded in several steps:

1. Familiarization and initial reading

Each guideline was read in full to gain an overview of its structure, tone, and explicit ethical commitments. During this stage, preliminary notes were taken on recurring themes such as justice, responsibility, cultural sensitivity, and evaluator competence.

Table 1. Coding scheme.

Concepts	Aristotle's definition	Analytical questions for guidelines	Possible indicators in the text
Telos (purpose/end)	Everything has a goal or purpose which determines its value.	What goal is formulated for the guidelines? Is it control, efficiency, responsibility, or flourishing?	Formulations about purpose, final goals, and what the evaluation "should lead to."
Eudaimonia (the good life, flourishing)	The highest good: the full development of human beings and society.	Do the guidelines promote human and societal flourishing, or technical utility?	References to human development, well-being, sustainability, and social benefit.
Arete (virtue/ excellence)	Good character develops through practice; virtues guide good actions.	Which virtues are reflected in the guidelines? Justice, courage, temperance, honesty?	Words or requirements related to integrity, justice, respect, objectivity, responsibility.
Phronesis (practical wisdom)	Practical judgment in complex situations; more than simply following rules.	Do the guidelines allow for professional judgment, or is ethics reduced or procedures?	Space for contextual adaptation, interpretation, reflection, and professional responsibility.
Mesotes (the golden mean)	Ethical action is a balance between extremes.	Is balance promoted between different demands (e.g. objectivity vs. respect, efficiency vs. justice)?	Text about trade-offs, proportionality, or balancing interests.
Zoon politikon (the human as a political animal)	Humans realize their full nature in community; the aim of politics is the common good.	Is evaluation positioned as a society-building and democratic practice, or as a technical tool?	References to society, democracy, participation, citizens, and the common good.

2. Deductive coding using Aristotelian concepts

The guidelines were then coded deductively using the six Aristotelian concepts as predefined categories. For each concept, we applied the analytical questions and indicators outlined in the coding scheme. For example, under *telos*, we examined how each organization formulates the overarching purpose of evaluation ethics; under *phronesis*, we assessed whether the guidelines encourage professional judgment or emphasize procedural compliance.

3. Comparative thematic analysis

After coding each document, we conducted a cross-case comparison to identify similarities, differences, and patterns. This step allowed us to detect broader tendencies, for instance, that "all guidelines underline responsibility and societal benefit as the overall aims (telos)" but differ in how strongly they emphasize social transformation versus professional quality.

4. Interpretation through Aristotelian theory

The final step involved interpreting the comparative findings through the lens of Aristotelian ethics. This interpretive layer enabled us to discuss how the guidelines relate to virtue development, practical wisdom, and the political dimension of evaluation. For example, the article notes that AfrEA and CES "encourage contextual sensitivity, dialogue, and ethical reflection, which can be considered as forms of making use of practical wisdom."

The analysis is hermeneutic rather than mechanical. While the coding scheme provides structure, the interpretation relies on close reading, contextual understanding, and theoretical reflection. This aligns with the Aristotelian emphasis on phronesis, ethical judgment sensitive to context, mirroring the very phenomenon under study.

Examined guidelines

This section provides a brief introduction to the organizations whose ethical guidelines are analyzed and outlines the development of their respective frameworks over time. For each organization, key historical context, institutional mandate, scope of activity, and significant milestones in the development and revision of their guidelines are highlighted. This background offers a foundation for understanding both the similarities and differences in the ethical principles discussed in the subsequent sections.

AfrEA's the African Evaluation Principles

AfrEA was established in 1999 in response to limited opportunities for professional networking, capacity building, and exchange among evaluators in Africa. At the time, evaluators often worked in isolation, and formal training in evaluation methodologies and standards was scarce, with evaluation practice largely carried out by technical specialists or management consultants (African Evaluation Association, n.d.). *The AfrEA Evaluation Principles* (AEP) were adopted in 2002 (Blaser Mapitsa and Polzer Ngwato, 2020) to

guide evaluation practice in service of the peoples of Africa and are intended for use by individuals, organizations, and networks involved in financing, commissioning, managing, conducting, and teaching evaluation across the continent. The AEP has undergone several revisions, most recently in 2021.

AEA's Guiding Principles

Two U.S.-based professional evaluation associations emerged in 1976: the Evaluation Network, composed primarily of university professors and school-based evaluators, and the Evaluation Research Society, which consisted mainly of government-based and university evaluators. In 1985, these two organizations merged to form the AEA (Preskill and Russ-Eft, 2005). The *Guiding Principles* constitute a foundational framework for sound and ethical evaluation practice. Originally developed in 1994, the principles have undergone broad review within the AEA membership and are periodically revised to reflect developments in the field, most recently in 2025. To support dissemination and application, the AEA Ethics Committee collaborated with practitioners in 2006–2007 to develop a training package designed for workshop-based instruction and professional development (American Evaluation Association, n.d.).

It should be noted that the AEA Guiding Principles were revised after the completion of our analysis. While the revised version places greater emphasis on context, stakeholder inclusion, positionality, equity, and the common good, the analysis reported here is based on the version that was in force during data collection and coding (Table 2).

AES' Guidelines for the Ethical Conduct of Evaluations

The AES officially commenced in 1987 (Sharp, 2003) and is a professional, member-based organization dedicated to advancing theory, practice, and use of evaluation across diverse sectors (Australasian Evaluation Society, n.d.). The original AES evaluation guidelines were endorsed by the Society's Board in December 1997 following extensive development and consultation with members. Since 1998, the guidelines have functioned as a key training resource and have been widely adopted by organizations to inform evaluation policy and practice. In December 2000, the guidelines were formally incorporated into the AES Code of Ethics, which is binding for all members. Revised versions of both the Code of Ethics and the Guidelines were subsequently endorsed in 2013.

CES' Guidance for Ethical Evaluation Practice

The CES, founded in 1981, is a national professional association that promotes the development and use of evaluation across diverse sectors. CES is particularly associated with program evaluation, notably through its publication, the *Canadian Journal of Program Evaluation*, but its scope encompasses evaluation theory and practice more broadly. CES supports professional development, networking, and the dissemination of high-quality evaluation practices among members from academia, government, private, and non-profit sectors (Canadian Evaluation Society, n.d.). CES' Guidance for Ethical Evaluation Practice was first published in 1995, but since then there has been considerable change in the practice of evaluation in Canada. The guidelines have undergone several changes; the latest revision was made in 2017.

UNEG's Ethical Guidelines for Evaluation

The UNEG is an interagency network of UN evaluation units, currently comprising 55 member agencies and 7 observers. Guided by the *UNEG Principles of Working Together* and its *2025–2032 Strategy*, UNEG aims to promote high-quality, independent evaluation to strengthen accountability, learning, and evidence-based decision-making across the UN system (United Nations Evaluation Group, n.d.). The UNEG *Ethical Guidelines for Evaluation* first published in 2008 and revised in 2020, provide a framework for professional conduct consistent with UN standards, core values, and humanitarian principles. They align with the UN Charter, Staff Regulations, and international norms, promoting integrity, professionalism, respect for diversity, and impartial, independent evaluation practices across the UN system.

Results and analysis

This section presents the empirical results of the comparative analysis. Drawing on the Aristotelian coding scheme, we examine how each set of ethical guidelines articulates its purpose, virtues, practical reasoning, and societal orientation. The comparative findings are summarized in Tables 3 to 6 and discussed in greater detail below. The analysis highlights both shared ethical commitments and notable divergences across the documents, which together illuminate the normative tensions at the center of contemporary evaluation ethics.

Empirical comparative analysis

The five guideline documents have many traits in common but also differences related to cultural and institutional contexts. All guidelines underline responsibility and societal benefit as

Table 2. American Evaluation Association's Guiding Principles (American Evaluation Association, 2018).

Concept	Analysis questions	Results
Telos	What goal is formulated for the guidelines? Is it control, efficiency, responsibility, or flourishing?	The purpose is to promote quality and ethical responsibility in evaluation, with the aim of contributing to the improvement of society.
Eudaimonia	Do the guidelines promote human and societal flourishing, or technical utility?	Promotes human and societal flourishing through principles of justice, respect, and social responsibility.
Arete	Which virtues are reflected in the guidelines? Justice, courage, temperance, honesty?	Highlights key virtues such as integrity, respect, justice, and competence as guiding ideals.
Phronesis	Do the guidelines allow for professional judgment, or is ethics reduced or procedures?	Acknowledges the evaluator's professional judgment and encourages ethical reflection in complex contexts.
Mesotes	Is balance promoted between different demands (e.g. objectivity vs respect, efficiency vs justice)?	Emphasizes balance between objectivity and respect, as well as between the interests of clients and participants.
Zoon politikon	Is evaluation positioned as a society-building and democratic practice, or as a technical tool?	Views evaluation as a contribution to democratic governance and collective learning processes.

Table 3. Australasian Evaluation Society's Guidelines for the Ethical Conduct of Evaluations (Australasian Evaluation Society, 2013).

Concept	Analysis questions	Results
Telos	What goal is formulated for the guidelines? Is it control, efficiency, responsibility, or flourishing?	The purpose is to ensure that evaluations are conducted ethically, with integrity and respect, to generate trustworthy knowledge and contribute to social improvement. The telos is responsibility and societal benefit rather than control or technical efficiency.
Eudaimonia	Do the guidelines promote human and societal flourishing, or technical utility?	The guidelines promote human and societal flourishing by emphasizing respect for people, equity, and social responsibility. They aim to enhance collective learning and well-being, not only methodological quality.
Arete	Which virtues are reflected in the guidelines? Justice, courage, temperance, honesty?	Central virtues include integrity, respect, justice, and responsibility. Evaluators are expected to demonstrate honesty, fairness, and competence, reflecting a virtue-ethical orientation.
Phronesis	Do the guidelines allow for professional judgment, or is ethics reduced or procedures?	The guidelines emphasize the use of professional judgment, contextual understanding, and ethical reflection. They acknowledge complexity and encourage evaluators to make reasoned, value-based decisions.
Mesotes	Is balance promoted between different demands (e.g. objectivity vs. respect, efficiency vs. justice)?	The guidelines promote balance between objectivity and empathy, between the needs of clients and the rights of participants, and between transparency and confidentiality, reflecting the "golden mean."
Zoon politikon	Is evaluation positioned as a society-building and democratic practice, or as a technical tool?	Evaluation is viewed as a collective, socially embedded practice that supports democratic participation, cultural respect, and shared learning, contributing to the common good.

the overall aims (telos). This shared orientation is visible in the wording of the documents themselves. The UNEG guidelines, for example, refer explicitly to beneficence and to doing good for people and the planet, while the AfrEA principles emphasize socially just change, local ownership, and African values. The AEA principles similarly stress social responsibility, integrity, and respect for people. AEA and AES focus mainly on quality, responsibility, and professional ethics, while CES and AfrEA extend the goals to comprehend social justice and cultural relevance as well. UNEG relates the aim to sustainable development goals and human rights. The most evident difference is the degree of societal orientation, ranging from professional quality to social transformation. None of the guidelines describe evaluation solely as a technical or controlling instrument. Eudaimonia, here interpreted as human welfare, is emphasized in all guidelines, but to varying degrees. AfrEA and CES highlight a holistic notion of flourishing, considering both human and ecological well-being. AEA and AES address this more indirectly through principles of justice and respect, while UNEG underscores both human and planetary flourishing through the concept of beneficence. All guidelines emphasize respect for people, justice, and responsibility as core values. This indicates a clear shift away from technical efficiency toward value-based evaluation. The most frequently recurring virtues (aretê) are integrity, justice, respect, and responsibility. These virtues are not merely

Table 4. Canadian Evaluation Society's (CES) Guidance for Ethical Evaluation Practice (Canadian Evaluation Society, 2017).

Concept	Analysis questions	Results
Telos	What goal is formulated for the guidelines? Is it control, efficiency, responsibility, or flourishing?	The goal is to guide evaluators toward ethically responsible practice grounded in shared values: respect for rights, truth, and accountability. The telos is to promote integrity and social good, not control or procedural compliance.
Eudaimonia	Do the guidelines promote human and societal flourishing, or technical utility?	The guidelines explicitly promote human and ecological well-being by emphasizing respect for people, communities, and the natural world. They connect evaluation to social justice and reconciliation, fostering holistic flourishing rather than mere methodological rigor.
Arete	Which virtues are reflected in the guidelines? Justice, courage, temperance, honesty?	Key virtues include integrity, honesty, respect, fairness, and accountability. The emphasis on relational ethics, truth-telling, and equity reflects a virtue-ethical orientation aimed at cultivating good evaluators rather than mere rule-followers.
Phronesis	Do the guidelines allow for professional judgment, or is ethics reduced or procedures?	Ethics is framed as value-based reflection and context-sensitive judgment. Evaluators are encouraged to deliberate ethically in complex situations and use practical wisdom to navigate conflicting values, strongly expressing *phronesis*.
Mesotes	Is balance promoted between different demands (e.g. objectivity vs respect, efficiency vs justice)?	The guidance seeks balance between truth and compassion, between client needs and community rights, and between methodological rigor and cultural sensitivity. It advocates proportionate, context-dependent judgment rather than rigid rules.
Zoon politikon	Is evaluation positioned as a society-building and democratic practice, or as a technical tool?	Evaluation is framed as a moral and social practice contributing to reconciliation, equity, and the public good. It highlights evaluators' roles as social actors responsible to communities and to the collective pursuit of justice and learning.

inferred through the Aristotelian framework but are explicitly articulated throughout the guidelines. Integrity is a recurring principle across AEA, AES, and UNEG, while the other guidelines place stronger emphasis on inclusiveness, cultural relevance, dignity, and accountability. AfrEA adds cultural humility and inclusiveness, while CES emphasizes honesty and accountability toward both people and nature. AEA and AES, on the other hand, focus more on professional competence and fairness. UNEG formulates its virtues in line with UN values, that is integrity, respect for diversity, and accountability. When it comes to practical wisdom (phronesis), the guidelines are quite similar, as they all pinpoint the importance of ethical judgment rather than strictly following the rules. This tendency is reflected in recurring references to professional discretion, contextual adaptation, and ethical reflection. Rather than presenting ethics as a matter of mechanical compliance, several of the guidelines explicitly acknowledge that evaluators must interpret principles in relation to specific circumstances, stakeholders, and value conflicts. Especially CES and AfrEA urge for the need of reflection and contextual adaptation, while AEA and AES describe how to behave correctly. UNEG

Table 5. The African Evaluation Association's the African Evaluation Principles (African Evaluation Association, 2021).

Concept	Analysis questions	Results
Telos	What goal is formulated for the guidelines? Is it control, efficiency, responsibility, or flourishing?	The stated goal is to ensure that evaluation practice in Africa is aligned with African societies' values, gives voice and power to African stakeholders, addresses power asymmetries, and supports socially just change. It is less about control or purely technical efficiency and more about empowerment and contextual relevance.
Eudaimonia	Do the guidelines promote human and societal flourishing, or technical utility?	The principles actively promote human and societal flourishing: They emphasize cultural relevance, indigenous knowledge, empowerment of communities, and redress of historical inequalities. The focus is on holistic flourishing rather than only utilitarian or technical goals.
Arete	Which virtues are reflected in the guidelines? Justice, courage, temperance, honesty?	Key virtues include respect, justice, fairness, inclusivity, cultural humility, and accountability. For example, one principle is "Ethically sound" which calls for respect for the dignity, knowledge and experience of individuals, communities, and nature. These reflect a virtue-ethical orientation.
Phronesis	Do the guidelines allow for professional judgment, or is ethics reduced or procedures?	The AEP emphasizes context, local knowledge, relationships, power dynamics and implies evaluators need to make wise, situational decisions. For instance, "Powerful for Africans" principle calls for processes that enable African actors to lead and exercise judgment. Thereby, they foster practical wisdom rather than only procedural compliance.
Mesotes	Is balance promoted between different demands (e.g. objectivity vs respect, efficiency vs justice)?	The principles call for balancing global/Western paradigms with African indigenous ones, balancing evaluator independence with local partnership, balancing rigor with relevance. For example, addressing power asymmetries while maintaining credibility and utility. This demonstrates a "golden mean" approach.
Zoon politikon	Is evaluation positioned as a society-building and democratic practice, or as a technical tool?	The AEP frames evaluation as a societal and democratic practice: enhancing ownership, strengthening domestic capacities, recognizing communal values (such as Ubuntu), and contributing to social justice and transformation. It is clearly more than a technical instrument.

stresses the need to balance different values. Mesotes, the golden mean, can be discerned in how the guidelines aim for a balance between conflicting demands. However, they have a slightly different focus. AEA and AES focus on balancing objectivity and respect, CES and AfrEA on balancing global and local perspectives, and UNEG on balancing transparency and confidentiality. Finally, all documents position evaluation as a societal practice (cf. Zoon politikon), but with varying emphasis. AEA and AES stress collective learning processes, CES and AfrEA consider evaluation a means for social change and justice, and UNEG relates evaluation to global responsibility and the common good. Especially CES, AfrEA, and UNEG describe evaluation as contributing to democracy, participation, and collective learning rather than merely as a management tool. In sum, the guidelines are fundamentally aligned around

Table 6. United Nations Evaluation Group's Ethical Guidelines for Evaluation (United Nations Evaluation Group, 2020).

Concept	Analysis questions	Results
Telos	What goal is formulated for the guidelines? Is it control, efficiency, responsibility, or flourishing?	The goal is to guide UN evaluation practice so that it is ethical, aligned with UN values (integrity, professionalism, and respect for diversity) and supports the 2030 Agenda. The telos emphasizes responsibility and ethical conduct rather than purely control or technical efficiency.
Eudaimonia	Do the guidelines promote human and societal flourishing, or technical utility?	The guidelines explicitly state principles such as beneficence ("striving to do good for people and planet while minimising harms"). This points toward flourishing, both human and environmental, not only methodological utility.
Arete	Which virtues are reflected in the guidelines? Justice, courage, temperance, honesty?	The virtues emphasized include integrity (honesty, truthfulness, independence, impartiality), accountability (being answerable); respect (for diversity, dignity); and beneficence (goodness, minimizing harm. So the guidelines embody a virtue-ethical orientation.
Phronesis	Do the guidelines allow for professional judgment, or is ethics reduced or procedures?	The document recognizes that evaluations occur in complex, dynamic contexts and encourages reflection and judgment (e.g. recognizing risks, power dynamics, context) rather than simply procedural checklists. Thus, it supports practical wisdom.
Mesotes	Is balance promoted between different demands (e.g. objectivity vs respect, efficiency vs justice)?	The guidelines aim for balance: for example independence and impartiality alongside respect and stakeholder engagement; minimizing harm while supporting systemic benefits; transparency while protecting confidentiality. In that sense, a "golden mean" is implicit.
Zoon politikon	Is evaluation positioned as a society-building and democratic practice, or as a technical tool?	The guidelines position evaluation as part of the UN's mission, supporting human rights, sustainable development and global good. Evaluation is framed as contributing to the public good, not merely a technical instrument.

the concepts of respect, justice, and responsibility but differ in the degree of societal and cultural embedding. These differences are grounded in how the documents themselves articulate ethical responsibilities, evaluator roles, stakeholder relationships, and societal purposes, even if the comparative interpretation is guided by the Aristotelian framework. AfrEA and CES are the most transformative and contextual, whereas AEA and AES are more professionally normative. These findings echo Schwandt's (2007) critique that ethical guidelines often overlook the institutional and cultural conditions that shape evaluators' work. He argues that ethics in evaluation cannot be confined to individual conduct but must also address the normative orientations embedded in social institutions and organized practices. Our results support this view in that the guidelines differ in how explicitly they address contextual influences: AfrEA, CES, and UNEG articulate these considerations more directly, whereas the other guidelines emphasize professional principles to a greater extent.

Theoretical analysis

For AEA and AES, the telos appears as ensuring good professional practice and credibility, which can be regarded as a form of technical and moral order. The other three guidelines, for CES, AfrEA, and UNEG, go further, and their telos comprehends contributing to human flourishing and societal development. These ambitions align with Aristotle's idea that every action should aim toward eudaimonia (the common good). According to Aristotle, "the good life" is achieved through practices that benefit the entire society, and not through rules. AfrEA and CES express this most clearly by relating ethics to reconciliation, justice, and ecological sustainability. UNEG's guidelines show similar ambitions, but on a global scale. AEA and AES, in contrast, stop at ensuring fair procedures. All guidelines are built on virtues (aretê) such as integrity, justice, and respect. Aristotle would, however, argue that virtues must be cultivated through habit and practical experience and not just through rules. This perspective is recalled by CES and AfrEA, which promote this by seeing the evaluator as a moral agent responsible for the community, whereas AEA and AES deal with virtue as more of a professional standard. Phronesis is central to Aristotle's ethics and illustrates the ability to act rightly in different situations. CES and AfrEA encourage contextual sensitivity, dialogue, and ethical reflection, which can be considered forms of making use of practice wisdom. AEA and AES show tendencies of phronesis but often remain at the level of general principles. UNEG emphasizes balancing rules and reflection, which can be seen as an attempt to institutionalize phronesis within a global framework. The golden mean (mesotes) is to be found in all guidelines, as everyone expresses the need for balance. Nevertheless, they do it in different ways: between independence (AEA, AES) and collaboration (CES, AfrEA), efficiency (AEA, AES) and justice (CES, AfrEA), and global (UNEG) and local (AfrEA, CES) considerations. This pursuit of balance reflects Aristotle's view of virtue as a mean between extremes, where sound judgment is decisive. For Aristotle, humans are social and political animals, and ethical life arises within community. AfrEA and CES articulate this most explicitly by seeing evaluation as a tool for participation and collective learning. UNEG expresses the same idea on a global scale as part of humanity's shared responsibility. AEA and AES show a weaker connection to the political and societal, emphasizing professional integrity instead.

From an Aristotelian perspective, the guidelines can be understood as different attempts to reconcile rule-governed ethics (ethics as norms) with virtue development (ethics as praxis). AfrEA and CES are closest to Aristotle's ideal of phronesis and eudaimonia, as they emphasize ethics as lived experience in community and strive toward societal flourishing. AEA and AES represent a more formal ethics, necessary but not sufficient for the good life in the Aristotelian meaning. UNEG occupies a middle ground, where global responsibility and balancing values give the guidelines a clear telos (aim) oriented toward the common good.

Discussion

The aim of this study was, from an Aristotelian ethical perspective, to compare and analyze a varied set of international and regional ethical guidelines for the evaluation field, with a focus on their underlying principles, modes of operationalization, and cultural contexts. The results show that the guidelines, to varying degrees, seek to balance universal ethical ideals with context-specific requirements, a balancing act that constitutes the core analytical problem of this study. By applying Aristotle's concepts of *telos*, *aretê*, *phronesis*, and *zoon politikon*, both

shared features and key differences emerge in how the organizations define the purpose of ethics, the role of the evaluator, and the relationship between evaluation, society, and culture. In doing so, this study highlights not only *what* the guidelines prescribe but also *what kind of ethical practice* they implicitly encourage. In relation to the aim, the results indicate that the guidelines collectively express an aspiration to formulate shared ethical principles, but that this aspiration is simultaneously constrained by different historical experiences, sociopolitical conditions, and value-based traditions. This confirms that the question of whether universal evaluation ethics is possible cannot be answered without considering the cultural and political dimensions that shape how ethics is understood and practiced in different parts of the world.

The results of this study can be understood in relation to the broader discussion on evaluation ethics, which several scholars have emphasized over the past two decades. As Simons (2006) and Morris (2011, 2015) note, the field has evolved from focusing on methodological standards to increasingly emphasizing ethics as an integrated and normative dimension of evaluation practice. All the guidelines in our material reflect this development by highlighting values such as integrity, respect, and responsibility as central to professional evaluation. However, our analysis shows that the guidelines operationalize these values in different ways depending on the cultural and institutional contexts, which confirms Schwandt's (2018) argument that evaluation ethics can never be fully detached from the sociopolitical contexts in which evaluations are carried out.

This finding is also consistent with the broader argument advanced by Van den Berg et al. (2022), who suggest that evaluation ethics should move beyond narrow concerns with compliance and avoidance of harm toward a more comprehensive consideration of what constitutes ethically responsible and socially valuable evaluation practice. Their emphasis on ethical reasoning, societal consequences, and evaluators' responsibilities resonates with key Aristotelian concerns regarding practical wisdom (phronesis), human flourishing (eudaimonia), and the relationship between professional conduct and the common good. This present study contributes to this discussion by demonstrating how different ethical guidelines operationalize these broader ethical ambitions in distinct ways despite sharing a broadly similar ethical core.

Against this broader development in evaluation ethics, a central finding of this study is the variation in the telos articulated in the guidelines. This relates directly to the literature on the purposes of evaluation. Vedung (2010, 2025) describes the history of evaluation as characterized by shifts between rationalist, instrumental ideals and more value-oriented and deliberative perspectives. The documents from AEA and AES still reflect much of a professional-rationalist ideal in which ethics serves to ensure quality and credibility. In contrast, the guidelines from AfrEA and CES present a more transformative view, in line with Hood et al. (2015), who argue that evaluations should acknowledge value pluralism, power dynamics, and local norms as central ethical dimensions. That AfrEA and CES adopt such a position can be understood in the light of the experiences of colonial structures, inequality, and cultural marginalization shared by both Africa and indigenous communities in Canada, conditions that Hood et al. (2015) and Aronsson and Hassnain (2021) highlight as decisive for context-sensitive ethics. This likely contributes to their emphasis on cultural relevance, justice, and the strengthening of local influence.

Several scholars have also emphasized the tension between universal principles and context-specific practice. Onyura et al. (2023) note that the ambition to create universal ethical guidelines always risks colliding with the local norms and values that shape evaluation

situations. Our study confirms this tension: While some guidelines seek broad applicability through universal principles (particularly AEA and UNEG), others stress the importance of local relevance, cultural grounding, and contextual interpretation (AfrEA and CES). This aligns with Aristotle's idea of *phronesis*, where good ethical action is not about following general rules but about interpreting them in the light of the characteristics of the concrete situation. Both CES and AfrEA articulate this explicitly by emphasizing that ethical judgment requires dialogue, reflection, and relational work—an idea that also resonates with Simons (2006), who portrays the evaluator as a moral actor in a complex social practice.

Furthermore, our analysis shows that the differences between the guidelines concern not only content but also the conception of the evaluator's role, an issue that Morris (2015) particularly highlights. AEA and AES emphasize professional integrity and competence, which reflects a view of the evaluator as a technically skilled expert. AfrEA, CES, and UNEG, on the other hand, emphasize the evaluator as a societal actor whose work has political and social consequences. This perspective is shared by Schwandt (2018), who argues that evaluation is always part of a broader societal context and therefore can never be a neutral or solely technical activity.

An important implication of these findings is that the differences between the guidelines should not be overstated. While the documents vary in how strongly they emphasize cultural relevance, societal transformation, professional responsibility, or global concerns, they are nevertheless built around a remarkably similar set of core ethical values. Respect, integrity, responsibility, justice, and concern for societal well-being are present across all five guidelines. The primary source of variation therefore appears not to be disagreement about ethical principles as such, but rather differences in how these principles are interpreted, prioritized, and operationalized within particular cultural, institutional, and political contexts. This observation has important consequences for the debate on universal evaluation ethics. If evaluation communities largely share fundamental ethical aspirations, the challenge is less about establishing entirely new universal principles and more about creating opportunities for collective reflection on how common principles should be understood and applied in different contexts. In this sense, dialogue becomes a practical mechanism for negotiating ethical meaning across settings while preserving a degree of shared normative orientation.

Taken together, this study shows that the guidelines, at an overarching level, move in the same direction as contemporary research on evaluation ethics: from a focus on standards and procedures toward a broader and more relational perspective in which power, context, and societal benefit are acknowledged to play a greater role. But the results also show that this development takes different forms in different regions, which strengthens the argument made by Hood et al. (2015) and Onyura et al. (2023) that universal evaluation ethics can never be understood without considering the cultural and political dimension.

Aristotle's ethical conceptual apparatus has served in this study as a framework for making these differences visible. In particular, the concepts of *telos*, *phronesis*, and *zoon politikon* illustrate that ethics in evaluation is not only about what evaluators *ought* to do but also about what kind of society evaluation contributes to shaping—a question that Schwandt (2007, 2018) and Simons (2006) also highlight as central to contemporary discussions on ethics. This connection reinforces the conclusion that guidelines must support not only correct procedure but also the development of ethical judgment and the competence to navigate value conflicts in situated social contexts.

As argued in the introduction, it is a difficult challenge to manage the tension between calls for universal evaluation ethics and at the same time being sensitive to vital contextual factors, such as cultural and political traditions. The findings presented here suggest that this tension should not be understood as a simple opposition between universalism and relativism. Rather, the guidelines indicate the existence of a shared ethical core combined with different understandings of how ethical principles should be translated into practice. The challenge is therefore not only to identify common values but also to develop ways of discussing and negotiating their practical implications across diverse contexts. Considering the empirical results of this article, underscoring the importance of cultural differences in the ethical guidelines, it seems difficult to find any single best strategy for trying to balance universal evaluation ethics with practical relevance. A top-down strategy to impose universal ethical standards risks becoming an ethnocentric project for dominant organizations in the West and far-reaching contextualization can relativize ethics to such an extent that ethical communication within the global evaluation community is undermined. We suggest that one way of proceeding is to try turning cultural differences into a fruitful resource for expanding conversations on evaluation ethics, in line with the ideas of Thomas Schwandt (2007). Ideally speaking, this could work as a bottom-up strategy, which at least to some extent could develop common ethical principles and codes of conduct and at the same time acknowledge some differences as appropriate and legitimate. Evaluation scholars and practitioners can surely learn from different social, cultural, and political conditions, which could benefit mutual ethical insights. The findings of this study suggest that dialogue is particularly important because the guidelines differ less in their commitment to core ethical values than in how those values are interpreted and operationalized. Dialogue therefore provides a means of negotiating the meaning of shared principles across different cultural and institutional contexts without abandoning the aspiration for a common ethical foundation. After all, dialogue and inclusion have for a long time been key values for the evaluation field.

Methodological reflections

This study is a qualitative, interpretive document analysis, and several limitations should be acknowledged. First, because the corpus consists solely of publicly issued ethical guidelines, the analysis captures organizational ideals rather than the lived moral practice of evaluation. We have primarily examined stated norms and professed virtues rather than the habituated dispositions that constitute excellence in action. This gap matters because virtue for Aristotle is not merely what is prescribed but what is cultivated through practice: through repeated choices that shape character. Consequently, our findings cannot speak to whether evaluators become more virtuous through these documents or exercise *phronesis* (practical wisdom) in real evaluative environments.

Second, although the selection of five prominent organizations' guidelines was intentionally broad, it nonetheless reflects the perspectives of well-established and well-resourced actors. This may result in underrepresentation of communities and associations where ethical reflection is embedded in different social practices or indigenous traditions. Our sample therefore reflects specific kinds of institutional communities and may not capture the full range of ways in which evaluative ethics are formed and sustained across diverse communal contexts.

Third, this study applies an Aristotelian analytic lens. This is a strength insofar as concepts like *telos*, *eudaimonia*, *aretê*, *mesotes*, and *phronesis* foreground purposes, flourishing,

character, balance, and judgment, dimensions that are often muted in rule-centric ethics. Yet it is also a limitation. Framing the analysis around these categories can render certain phenomena, especially legible (e.g. how guidelines balance conflicting goods), while making other ethical grammars, such as decolonial, feminist, rights-based, or care-ethical vocabularies, less salient. Put differently, our interpretive grid highlights Aristotelian virtues but may underilluminate forms of moral reasoning that do not readily map onto *mesotes* (the mean) or character-centered excellence.

Finally, a document-based comparison infers purposes and conceptions of the good from textual formulations. While such inference is defensible in a virtue-ethical analysis, it should be complemented in future research by empirical work that traces how these guidelines shape habituation (training, mentoring, communities of practice), how evaluators learn to deliberate well, and whether organizational conditions permit the exercise of *phronesis*. Without such practice-facing evidence, the step from written norms to cultivated virtue, and ultimately to *eudaimonia* in evaluative practice and its effects, remains suggestive rather than demonstrated.

ORCID iD

Anna Petersén https://orcid.org/0000-0001-5675-9428

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of conflicting interests

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

References

African Evaluation Association (n.d.) *Our History*. African Evaluation Association. Available at: https://afrea.org/our-history/ (accessed 18 November 2025).

African Evaluation Association (2021) *African Evaluation Principles*. African Evaluation Association. Available at: https://afrea.org/ (accessed 21 November 2025).

American Evaluation Association (n.d.) *AEA Guiding Principles Training Package*. American Evaluation Association. Available at: https://www.eval.org/Evaluation-Policy/Policy-Statements/Guiding-Principles-Training-Package?utm_source=chatgpt.com (accessed 18 November 2025).

American Evaluation Association (2018) *Guiding Principles for Evaluators*. American Evaluation Association. Available at: https://www.eval.org/ (accessed 21 November 2025).

Aristotle (2012) *Nicomachean Ethics*. Chicago University Press.

Aronsson I-L and Hassnain H (2021) Evaluation and ethics in contexts of fragility, conflict and violence. In: van den Berg RD, Hawkins P and Stame N (eds) *Ethics for Evaluation: Beyond "Doing No Harm" to "Tackling Bad" and "Doing Good"*. Routledge, pp.83–104.

Australasian Evaluation Society (n.d.) *About the Australasian Evaluation Society*. Australasian Evaluation Society. Available at: https://www.aes.asn.au/about-us (accessed 19 November 2025).

Australasian Evaluation Society (2013) *Guidelines for the Ethical Conduct of Evaluations*. Australasian Evaluation Society. Available at: https://www.aes.asn.au/ (accessed 21 November 2025).

Canadian Evaluation Society (n.d.) *About CES*. Canadian Evaluation Society. Available at: https://evaluationcanada.ca/about/who-we-are/index.html (accessed 20 November 2025).

Canadian Evaluation Society (2017) *Guidance for Ethical Evaluation Practice*. Canadian Evaluation Society. Available at: https://evaluationcanada.ca/ (accessed 21 November 2025).

Dahler-Larsen P (2013) Evaluation as a situational or a universal good? Why evaluability assessment for evaluation systems is a good idea, what it might look like in practice, and why it is not fashionable. *Scandinavian Journal of Public Administration* 16(3): 29–46.

Evrigenis ID (1999) The doctrine of the mean in Aristotle's ethical and political theory. *History of Political Thought* 20(3): 393–416.

Field JC and MacIntyre Latta M (2001) What constitutes becoming experienced in teaching and learning? *Teaching and Teacher Education* 17(8): 885–895.

Fleischacker S (1999) From cultural diversity to universal ethics. Three models. *Cultural Dynamics* 11(1): 105–128.

Hood S, Hopson RK and Kirkhart K (2015) Culturally responsive evaluation. In: Newcomer KE, Hatry HP and Wholey JS (eds) *Handbook of Practical Program Evaluation*. John Wiley & Sons, pp.281–317.

IDEAS (2014) *IDEAS Code of Ethics*. International Development Evaluation Association. Available at: https://ideas-global.org/ (accessed 24 February 2026).

Kuçuradi I (2014) Approaches to ethics, approaches in ethics and the idea of a "universal ethics". In: Fløistad G (ed.) *Ethics or Moral Philosophy, Contemporary Philosophy: A New Survey* (vol. 11). Springer, pp.99–108.

Lang AF Jr (2020) Constructing universal values? A practical approach. *Ethics & International Affairs* 34(3): 267–277.

MacIntyre A (2007 [1981]) *After Virtue*. 3rd ed. University of Notre Dame Press.

Mapitsa CB and Ngwato TP (2020) Rooting evaluation guidelines in relational ethics: Lessons from Africa. *American Journal of Evaluation* 41(3): 404–419.

Mark MM, Eyssell KM and Campbell B (1999) The ethics of data collection and analysis. *New Directions for Evaluation* 1999: 47–56.

Morris M (2011) The good, the bad, and the evaluator: 25 years of AJE ethics. *American Journal of Evaluation* 32(1): 134–151.

Morris M (2015) Research on evaluation ethics: Reflections and an agenda. *New Directions for Evaluation* 148: 31–42.

Msoroka MS and Amundsen D (2018) One size fits not quite all: Universal research ethics with diversity. *Research Ethics* 14(3): 1–17.

Olsson J and Petersén A (2025) Utvärdering ur ett etiskt perspektiv. In: Norén Bretzer Y and Lindgren L (eds) *Utvärdering i en osäker tid: Komplexitet, problem och utveckling*. Studentlitteratur, pp.157–174.

Onyura B, Main E, Barned C, et al. (2023) The "what" and "why" of (un) ethical evaluation practice: A meta-narrative review and ethical awareness framework. *Canadian Journal of Program Evaluation* 38(2): 265–312.

Pagan NA (2008) Configuring the moral self: Aristotle and Dewey. *Foundations of Science* 13: 239–250.

Petersén A and Olsson J (2015) Calling evidence-based practice into question: Acknowledging phronetic knowledge in social work. *British Journal of Social Work* 45(5): 1581–1597.

Preskill H and Russ-Eft D (2005) *Building Evaluation Capacity*. Sage.

Schwandt TA (2007) Expanding the conversation on evaluation ethics. *Evaluation and Program Planning* 30: 400–403.

Schwandt TA (2018) Acting together in determining value: A professional ethical responsibility of evaluators. *Evaluation* 24(3): 306–317.

Sharp CA (2003) The historical perspectives on evaluation project. *Evaluation Journal of Australasia* 3(2): 4–5.

Simons H (2006) Ethics in evaluation. In: Shaw I, Green J and Mark M (eds) *The Sage Handbook of Evaluation*. Sage, pp.243–265.
Tsang NM (2007) Reflection as dialogue. *British Journal of Social Work* 37(4): 681–694.
United Nations Evaluation Group (n.d.) *History*. United Nations Evaluation Group. Available at: www.unevaluation.org/about/overview (accessed 28 November 2025).
United Nations Evaluation Group (2020) *Ethical Guidelines for Evaluation*. UNEG. Available at: https://www.unevaluation.org/ (accessed 21 November 2025).
Van den Berg RD, Hawkins P and Stame N (2022) *Ethics for Evaluation. Beyond "Doing No Harm" to "Tackling Bad" and "Doing Good"*. Routledge.
Vedung E (2010) Four waves of evaluation diffusion. *Evaluation* 16(3): 263–277.
Vedung E (2025) Sex vågor av utvärdering över atlantiska världen. In: Norén Bretzer Y and Lindgren L (eds) *Utvärdering i en osäker tid: Komplexitet, problem och utveckling*. Studentlitteratur, pp.33–53.

Author biographies

Anna Petersén is Senior Lecturer in Social Work at Örebro University. Her research focuses on evaluation, professional practice, and the organizational and institutional conditions surrounding human service organizations.

Jan Olsson is Professor of Political Science at Örebro University. His research examines policy processes, governance, evaluation, democracy, and institutional change, with particular interest in multi-level and cross-sector policymaking.