Evaluation of Interdisciplinary and Transdisciplinary Research
A Literature Review

Julie T. Klein, PhD

Abstract: Interdisciplinarity has become a widespread mantra for research, accompanied by a growing body of publications. Evaluation, however, remains one of the least-understood aspects. This review of interdisciplinary and transdisciplinary research evaluation categorizes lessons from the emergent international literature on the topic reviewed in 2007. It defines parallels between research performance and evaluation, presents seven generic principles for evaluation, and reflects in the conclusion on changing connotations of the underlying concepts of *discipline, peer,* and *measurement.* Interdisciplinary and transdisciplinary research performance and evaluation are both generative processes of harvesting, capitalizing, and leveraging multiple expertise. Individual standards must be calibrated, and tensions among different disciplinary, professional, and interdisciplinary approaches carefully managed in balancing acts that require negotiation and compromise. Readiness levels are strengthened by antecedent conditions that are flexible enough to allow multiple pathways of integration and collaboration. In both cases, as well, new epistemic communities must be constructed and new cultures of evidence produced. The multidisciplinary–interdisciplinary–transdisciplinary research environment spans a wide range of contexts. Yet seven generic principles provide a coherent framework for thinking about evaluation: (1) variability of goals; (2) variability of criteria and indicators; (3) leveraging of integration; (4) interaction of social and cognitive factors in collaboration; (5) management, leadership, and coaching; (6) iteration in a comprehensive and transparent system; and (7) effectiveness and impact.
(Am J Prev Med 2008;35(2S):S116–S123) © 2008 American Journal of Preventive Medicine

Introduction

Interdisciplinarity has become a widespread mantra for research, accompanied by a growing body of publications. Evaluation, however, remains one of the least-understood aspects. In the past, discussions of interdisciplinary and transdisciplinary evaluation did not constitute an identifiable literature. They were scattered across multiple forums, and they were longer on anecdotal, intuitive, and normative perspectives than on empirical, longitudinal, and large-scale studies. In the absence of clear guidelines, Laudel and Origgi[1] recount, faculty and administrators had to "muddle through." The three clusters of work in Figure 1,[1–28] though, form an emergent international literature identified in 2007 by cross-referencing publication citations, significant addresses, and discussions in electronic networks focused on the topic. Cluster 1 spans an international body of studies recognized in the April 2006 benchmark issue of *Research Evaluation* on inter-

disciplinary research assessment.[2] Cluster 2 centers on the concept of transdisciplinary team science in the U.S. highlighted in this supplement to the *American Journal of Preventive Medicine.*[14] Cluster 3 encompasses studies from the European transdisciplinary movement for trans-sector, problem-oriented research involving the participation of stakeholders in society.

The contexts of interdisciplinary and transdisciplinary research vary greatly, as well as the attendant methodologies and conceptual frameworks. Yet cross-cutting themes provide a comparative framework for thinking about evaluation that draws insights from qualitative and quantitative studies. This review defines parallels between research performance and evaluation, and then presents seven generic principles for evaluation. The conclusion addresses implications for the underlying concepts of *discipline, peer,* and *measurement.* Interdisciplinary and transdisciplinary research performance and evaluation are both generative processes of harvesting, capitalizing, and leveraging multiple kinds of expertise. Individual standards must be calibrated and tensions among different approaches carefully managed in balancing acts that require negotiation and compromise. Readiness levels are strength-

From the Interdisciplinary Studies Program/Department of English, Wayne State University, Detroit, Michigan

Address correspondence and reprint requests to: Julie T. Klein, PhD, 111 Linden Court, Ypsilanti MI 48197. E-mail: julietklein@comcast.net.
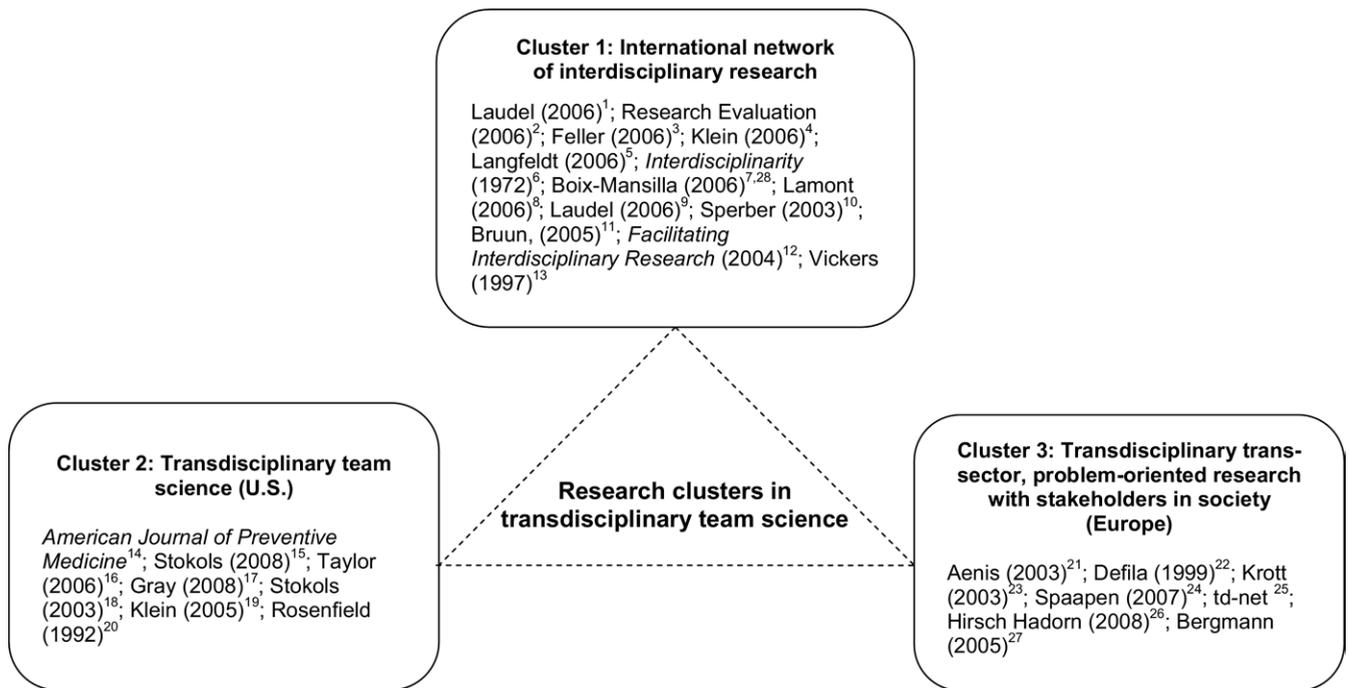
Figure 1. Clusters of emergent literature

ened by antecedent conditions that are flexible enough to allow multiple pathways of integration and collaboration. Appropriate epistemic communities must also be constructed and new cultures of evidence produced.

Research in the multidisciplinary–interdisciplinary–transdisciplinary environment is not a set of mutually exclusive categories. Research is too complex, Spaapen et al.[24] advise, to be put into boxes that ignore the particularities of context. In their introduction to this supplement, Stokols et al.[15] present recognized distinctions between multidisciplinary juxtapositions of disciplinary approaches and more robust interdisciplinary integrations and collaborations. In defining *transdisciplinary*, they adopt Rosenfield's connotation[20] of a process in which members of different fields work together over extended periods to develop novel conceptual and methodologic frameworks with the potential to produce transcendent theoretical approaches. This connotation is consistent with the earliest definition of *transdisciplinary*[6] as a common axiom that transcends separate disciplinary perspectives, exemplified by the overarching syntheses of general systems and ecology. A second major connotation in the European transdisciplinary movement should also be acknowledged: trans-sector, problem-oriented research involving a wider range of stakeholders in society. Both connotations are necessary for a full understanding of the spectrum of interdisciplinary and transdisciplinary research.

The evaluation of interdisciplinary and transdisciplinary research is a complex task. More than one discipline, profession, and field—or perhaps all three—are involved. Levels and subsystems differ, ranging from small

projects to national research systems, from the personal and interpersonal to organizational and systemic scales, and from academic settings to trans-sector projects with external stakeholders. Criteria also vary across stages, from ex ante to ex post assessments, and programs and projects differ by knowledge domain, institutional location, goals, and type of integration. The scope of integration, in turn, varies from middle-range and narrow-gauged or horizontal forms of interdisciplinarity among neighboring disciplines with compatible epistemologies to broad-gauged, vertical, and grand-scale forms among disciplines with more divergent epistemologies.[16,18] In short, as Feller[3] emphasized in a 2006 symposium on interdisciplinary research evaluation at the American Association for the Advancement of Science (AAAS), the reality of interdisciplinary evaluation is shaped by multiples: multiple actors making multiple decisions in varied organizational settings with context-dependent measures of quality. As a result, Spaapen et al.[24] add, quality is a relative concept determined by relations within the environment of a group and their goals. Research must "attune a pluralism of interests and values" within a dynamic set of programs and contexts and with a variegated group of stakeholders.[24]

The heterogeneity of the multidisciplinary–interdisciplinary–transdisciplinary environment defies the quest for a single best procedure for research performance or evaluation. Yet the emergent literature,[1–5,7–28] suggests seven generic principles of evaluation (Table 1): (1) variability of goals; (2) variability of criteria and indicators; (3) leveraging of integration; (4) interaction of social and cognitive factors in collaboration; (5) manage-

Table 1. Correlation of principles and references

Principle number	Evaluation principles
1	Variability of goals[5,7,11]
2	Variability of criteria and indicators[7,12,28]
3	Leveraging of integration[7,14–16,18,19,22,23,28]
4	Interactions of social and cognitive factors in collaboration[16,18,21,22,24,27,28]
5	Management, leadership, and coaching[7,9,11,17,19,22,24]
6	Iteration in a comprehensive and transparent system[16,18,21,24]
7	Effectiveness and impact[12,22,23,28]

ment, leadership, and coaching; (6) iteration in a comprehensive and transparent system; and (7) effectiveness and impact. Klein[4] defined these principles earlier, but they are placed here within an expanded comparative framework that incorporates new work.

Principle #1. Variability of Goals

Interdisciplinary and transdisciplinary research are not driven by a single goal. Based on a comparative analysis of evaluation procedures in Europe and the U.S., Langfeldt[5] concluded that sensitivity to context and flexibility are fundamental. Two studies[7,11] in Cluster 1 underscore the principle of variability. When an Academy of Finland integrative research team examined how well the Academy was accommodating interdisciplinary research in all funding categories based on the analysis of research proposals and interviews,[11] the most important reason cited for selecting an interdisciplinary approach was typically an epistemological goal: the production of new and broad knowledge of a particular phenomenon. Informants also cited new approaches that are interesting and hold potential as well as synergies stimulated by sharing knowledge, skills, or resources. Others mentioned the development of technical equipment or products such as information technology protocols, medicines, and measuring devices. Broadly speaking, methodological interdisciplinarity dominated over more-challenging conceptual and theoretical forms, achieved typically by combining concrete methods or research strategies from different fields in order to test a hypothesis, answer a research question, or develop a theory.

A similar variety of goals appeared when a team from the Interdisciplinary Studies Project at Harvard University[7] interviewed researchers in five organizations with extensive experience in conducting interdisciplinary research. In a project involving physicists assessing their mathematical theories of innovation and network behavior, researchers favored qualities such as "the ability to predict" unstudied social and biological phenomena and "tangible success" in explaining something that had not been explained previously. In a project com-

bining physiology, molecular biology, nanophysics, and materials science, scientists valued creation of an "unprecedented entity": for example, a vascularized artificial liver that "works" and has a "transforming effect" on organ transplantation surgical practice. Researchers engaged in pragmatic problem solving and product development placed a higher premium on viability, workability, and impact, while contributions seeking algorithmic models of complex phenomena were associated with simplicity, predictive power, and parsimony. Contributions seeking a more-grounded understanding of multidimensional phenomena, such as lactose intolerance or organ donation, favored work reaching new levels of comprehensiveness, careful description, and empirical grounding.[7] The key implication of this study is that variability of goals in turn drives variability of criteria and indicators of quality.

Principle #2. Variability of Criteria and Indicators

The Harvard team[7] identified two approaches to the assessment of interdisciplinary quality based on interview results. The first—conventional metrics—has been privileged traditionally. Informants reported being judged typically on indirect or field-based quality indicators: numbers of patents, publications, and citations; prestige rankings; and the approval of peers and a broader community. Hence, the first epistemic criterion in the study was consistency with multiple-"antecedent disciplinary knowledge." Credibility was strengthened by "fit" with disciplinary antecedents. Yet when work violated fundamental tenets or revealed limitations, additional justification was required.[7] Field-based measures, informants indicated, sidestep the question of what constitutes warranted interdisciplinary knowledge by relying on the social procedures of peer review, inter-subjective agreement, and consensus on what constitutes acceptable results. Informants were often critical of such "proxy" criteria, believing that they represent a strictly disciplinary assessment. More primary or epistemic measures of "good" work are needed that address the substance and constitution of the research, such as experimental rigor, aesthetic quality, fit between framework and data, and the power to address previously unsolved questions in a discipline.[28]

Other studies[12] affirm the principle of variability. The 2004 report *Facilitating Interdisciplinary Research*[12] from the U.S. National Academies of Science (NAS) cites outcomes in and feedback to multiple fields or disciplines; expanded expertise, vocabularies, and tool sets; the ability to work in more than one discipline; a greater proclivity toward interdisciplinary and transdisciplinary collaboration; and a widened sphere of professional reading. Individuals responding to national surveys preliminary to the report also cited participation in new subfields and departments as well as multidisciplinary advisory or review groups; new formal

affiliations; and the co-mentoring of doctoral students. Changing career trajectories were gauged by new appointments, recognition within and outside a person's original field, and, in areas such as sustainability and health outcomes, new public-policy initiatives and altered protocols in health management.[12]

Principle #3. Leveraging of Integration

Studies of interdisciplinary and transdisciplinary research call attention not only to outcomes but also to the quality of the process. Integration is widely considered the crux of interdisciplinarity,[29] and Krott[23] deems integration the critical point for evaluation in transdisciplinary projects. Likewise, the Harvard Project[7] highlighted the epistemic criterion of balance in weaving perspectives into a coherent whole, and integration was one of four "hot spots" identified in the 2006 AAAS symposium, in the form of "reaching effective syntheses."[7] The heart of the process, Boix-Mansilla[28] explains, is leveraging integration. In linking processes of intellectual integration and collaboration, the introduction to this supplement[15] and studies[18] of the Transdisciplinary Tobacco Use Research Centers (TTURCs) also stress the role of antecedent conditions, including frequent opportunities for communication, structural support, and a transdisciplinary ethic.

Two sets of guidelines[19,22] stress the importance of engaging integration from the beginning. Klein's "Guiding Questions for Integration"[19] was created for ex ante evaluation of grant proposals in the TTURCs program and subsequently revised for Land & Water Australia's key document on integration in natural resource management. Klein highlights a number of evaluation questions aimed at fostering integration and monitoring relationships among organizational, methodologic, and epistemologic components of a project or program. Is the spectrum of disciplines and fields too narrow or too broad for the task at hand? Have relevant approaches, tools, and partners been identified? Is the structure flexible enough to allow for shifting groupings of individuals and context-related adaptations, deletions, and additions? Has synthesis unfolded through patterning and testing the relatedness of materials, ideas, and methods? Have known integrative techniques been utilized, such as the Delphi method, scenario building, general systems theory, and computer analyses of stakeholders' perspectives? And, is there a unifying principle, theory, or set of questions that provides coherence, unity, or both?

Defila and DiGiulio's[22] catalogue of criteria emerged from a study of trans-sector transdisciplinary research commissioned by the Swiss National Science Foundation. The catalogue provides a comprehensive set of building blocks to help construct either a self-evaluation or an external evaluation of a research program. The power of the generative approach to evaluation lies in its flexibility. All categories in the catalogue of criteria may not apply at all phases (e.g., scientific quality or integration/synthesis or project organization/management). The timing and number of evaluations can also be adjusted throughout stages, and the questions of who performs the evaluation and the weighting of criteria are left open, too.

Principle #4. Interaction of Social and Cognitive Factors in Collaboration

The studies of transdisciplinary collaboration in Clusters 2 and 3 (Figure 1) emphasize the interaction of social and cognitive factors. While recognizing familiar indicators such as publications, the logic model that emerged from studies of the TTURCs accords greater weight to collaboration and does not sharply separate cognitive–epistemic and social factors.[16,18] Comparably, Spaapen et al.[24] describe research in the multidisciplinary–interdisciplinary–transdisciplinary environment as a "social process of knowledge production." Studies of interdisciplinary collaboration concur (Amey and Brown,[30] Derry et al.[31]). In Cluster 1 (Figure 1), Boix-Mansilla[28] highlights the need to calibrate separate standards while managing tensions through compromise and negotiation. The ongoing and systematic communication of research partners and subprojects lessens the likelihood of shortfalls of integration. The clarification and negotiation of differences lessen misunderstanding and strengthen the conditions for consensual modes of work. Intellectual integration is leveraged socially through mutual learning and joint activities that foster common conceptions of a project or program and common assessments. Mutual knowledge emerges as novel insights are generated, disciplinary relationships redefined, and integrative frameworks built. Within a heterogeneous mix of disciplines, though, compromises must be made, and the best option may be a partial, negotiated consensus.

Drawing on experiences in trans-sector transdisciplinarity within European landscape studies, Aenis and Nagel[21] formulated two axiomatic considerations for evaluation: the meta-level of interdisciplinarity (communication among researchers) and participation (communication between researchers and regional actors). Communication and negotiation also lie at the heart of the *Evalunet Guide for Formative Evaluation of Research Projects*,[27] an initiative of the Institute for Social–Ecological Research in Germany. The question-based guide provides both basic and detailed criteria based on the empirical study of projects in European research institutes. Evaluation is defined a collaborative and discursive learning process. Individuals first address questions by themselves, and then arrive at a common plan together, rather than imposing a priori a universal scoring method. Like the Defila and DiGiulio catalogue,[22] the detailed criteria of the *Evalunet* guide are also flexible.

Principle #5. Management and Coaching

Competence, Klein[19] and Defila and DiGiulio[22] also concur, is defined partly in terms of how well the management of projects and programs implements consensus building and integration. Therefore, evaluation must consider how well the organizational structure fosters communication, including networking among subprojects. The organizational chart and task distribution must allow time for interaction, joint work activities, common instruments, and shared decision making. If a group is pushed too quickly toward integration, the crucial activities of building rapport and exploring ways to understand how each discipline approaches a research question are shortchanged, ultimately shortchanging the quality of the integration. Comparably, as participants[7] in the 2006 AAAS symposium exhorted, in the peer-review process expertise must be carefully managed if panelists are to calibrate their individual beliefs about the meaning of quality.

Leadership is another prominent theme. Gray[17] in this supplement categorizes three types of leadership tasks for transdisciplinary research. **Cognitive tasks** focus on meaning making through a mental model or mindset. Visioning and reframing stimulate ideas about how disciplines might overlap in constructive ways that generate new understandings and encourage collaborative work modes. **Structural tasks** entail management issues of coordination and information exchange, including focus and defining objectives, recruitment of expertise, and accountability for deadlines and deliverables. External boundaries must be spanned, and internal linkages and information flows brokered across different disciplinary cultures, status hierarchies, and organizational structures. **Process tasks** ensure constructive and productive interactions among team members, with the attendant subtasks of designing meetings, determining ground rules, identifying tasks that move partners toward their objectives, building trust, and ensuring effective communication (and, if necessary, removing a member). Ultimately, Gray[17] conceptualizes transdisciplinary collaboration as innovation networks, underscoring the need for network stability, knowledge mobility, and innovation appropriability.

Recently, the theme of coaching both the research and evaluation processes has emerged in Clusters 2 and 3 (Figure 1). Klein[19] and Defila and DiGiulio[22] recommend also using their evaluation guidelines to nurture integration during the actual course of research. Spaapen et al.[24] describe their Research Embedment and Performance Profile (REPP), which emerged from studies of agricultural and pharmaceutical research, as a coaching model rather than a jury model. The REPP facilitates the graphic depiction of the main activities of a group (e.g., publications, collaboration, innovation) and its performance, fostering self-reflection about process, performance, and mission.[24] For peer review,

the Academy of Finland integrative research team[11] recommends that national funding agencies coach the interdisciplinary and transdisciplinary process, and Laudel[9] cites an exemplary model. The German *Sonderforschungsbereiche* (SFBs) are networks of research groups that receive funding for collaborative research programs. The core of the review process is a series of group discussions among the reviewers and between reviewers and applicants. A group or center is also evaluated every third year by largely the same reviewers. Repeating the process ensures that reviewers gain the necessary competence and a communication base over time, facilitated by the empowerment of applicants and the enforced interdisciplinary learning of reviewers.[9]

Principle #6. Iteration and Transparency in a Comprehensive System

Studies of interdisciplinary and transdisciplinary collaboration highlight the overriding importance of iteration to ensure collaborative input, transparency, and common stakeholding. In the TTURCs logic model,[16,18] indicators are not restricted to a single phrase. They have a feedback relationship that a strictly linear model of evaluation cannot capture. The logic model moves from the basic activities of centers (training, collaboration, and integration) and the earliest expected outcomes. Basic activities lead to new and improved methods, science, and models that are tested and lead to publications. Publications, in turn, foster recognition and the institutionalization of transdisciplinary research that feed back on the overall infrastructure and capacity of centers, resulting in increased support for basic activities. They also provide a content base for communicating results to a broader community. Recognition, in turn, provides a secondary impetus for communications and publications. Policy implications result as well from communications and publications, while translation to practice is influenced by improved interventions. Health outcomes, for example, are influenced both by treatments and health practices related to policy changes.[16,18]

Two models in Cluster 3 furnish insights from fields of application. Aenis and Nagel[21] used logical-framework (log-frame) analysis to define impact indicators in agricultural research, based on the systematic elaboration of objectives at the beginning. The central insight is that the mobility of participants and interaction and communication patterns furnish a heuristic for identifying differences in social domains or contexts for knowledge production. In each context, differing expectations exist, with attendant norms, values, and priorities.[21] The REPP method of Spaapen et al.[24] facilitates the reconstruction of both the relevant environment and the performance of a group within it, seeking patterns and profiles rather than imposing a

priori measurements. A quantifiable benchmark, though, can be set for each indicator in consultation with researchers and policymakers. Scores are plotted on a radar-like graph that represents variegated activities. If a group claims to contribute to the development of sustainable greenhouse production, for example, the profile should show that empirically. The key dynamics are feedback to the mission of a program and transparency of criteria. Feedback allows for context-related adaptations that improve the research process and conceptual framework. Transparency requires that both evaluators and participants are informed of criteria from the outset and, ideally, are involved in defining them.[24]

Principle #7: Effectiveness and Impact

Principle #7 returns full circle to Principles #1 and #2: variability of goals drives variability of criteria and indicators. The third criteria of quality in the Harvard study was effectiveness in advancing epistemological understanding or pragmatic viability in concrete settings. Unintended consequences and unforeseeable long-term impacts, though, cannot be captured by a priori measures, and they may have multiple consequences. "Interdisciplinary impacts," Boix-Mansilla cautions, "are often diffused, delayed in time, and dispersed across diverse areas of study and patterns of citation practice."[28] Defila and DiGiulio agree, admonishing that many long-term effects cannot be predicted or checked in five-year periods, let alone annual measures.[22] In trans-sector transdisciplinary, Krott notes, different target groups also make use of knowledge in ways unknown at the start of a project.[23] Likewise, studies[16,18] of the TTURCs stipulate that the appropriate time frame for assessing returns on investment or the value-added contributions of large-scale transdisciplinary collaboration may require broad historical perspectives spanning two or more decades.

The NAS report *Facilitating Interdisciplinary Research*[12] cites numerous examples of long-term impacts that could not be predicted or measured fully at the outset. Research on nitrate and sulphate cycles, for instance, proved relevant not only for agricultural production but also for research on global climate change and the greenhouse effect. Developing the engineering technologies necessary to achieve space flight led to advances in computer control of engineering processes that subsequently fostered improvements in the reliability of industrial products and processes. Large programs also stimulate new understanding in multiple fields, a long-term effect evident in the Human Genome Project, the Manhattan Project, and in broad efforts such as the theory of plate tectonics and the development of the fiber-optic cable. Moreover, generative technologies such as magnetic resonance imaging are enhancing research capabilities in an expanding number of areas through new instrumentation and informational analysis.[12]

Conclusion: The Logic of *Discipline, Peer,* and *Measurement*

An emergent literature is a benchmark of both what is known and what remains to be known. Key insights from this literature appear in Table 2. Yet findings are still dispersed across multiple forums, even with systematic efforts to disseminate information by groups such as the Europe-based td-net.[25,26] Longitudinal empirical studies of interdisciplinary and transdisciplinary evaluation remain few in number and need testing in local contexts. Access to in vivo deliberations is still limited in peer review, and governments lack clearly defined and tested criteria for prioritizing funding across the spectrum of disciplinary and multidisciplinary–interdisciplinary–transdisciplinary research. And, more broadly, unquestioned assumptions about three underlying concepts—*discipline, peer,* and *measurement*—continue to cloud the discourse on evaluation.

Disciplines provide crucial knowledge, methodologies, and tools for interdisciplinary and transdisciplinary work. However, in many discussions, disciplines are still treated uncritically as monolithic constructs. Studies of disciplinarity reveal that disciplines exhibit a striking heterogeneity, and that boundary crossing has become a marked feature of contemporary research. Some disciplines, Vickers[13] observes, have undergone so much change that characterizing them as stable matrices with consensual evidentiary protocols is problematic. Some new interdisciplinary and transdisciplinary fields also reject disciplinarity in whole or in part, and, Sperber[10] observed in an online virtual seminar, the purpose of interdisciplinary work may aim to undermine current understanding in disciplines. A standard assessment procedure can help in charting a program's interactions within a broader environment and ensuring that work is sound and reliable.[24] Yet stringent evaluation criteria for both research and evaluation may be counterproductive, especially, Langfeldt[5] warns, for risk taking and "radical interdisciplinarity." Conflicting assumptions about quality meet head-on during *peer* review, whether in ex ante evaluations of grant proposals and priority setting in national research systems or in ex post assessments of research performance and outcomes. A "commonly agreed yardstick" must be developed to "moderate the conservative forces" of traditional research communities, safeguarding against bias.[5]

Identifying experts who fit the "problem space" is crucial, because they form an appropriate interdisciplinary epistemic community. The task is more difficult, though, in emerging fields where the criteria of excellence are not defined yet and the pool of qualified

Table 2. Key insights

Principle number	Evaluation principles	Key insights
1	Variability of goals	**Variances:** size, scope, scale, level and subsystem, degree of integration in multidisciplinary–interdisciplinary–transdisciplinary environment **Multiple goals:** for example, epistemologic or methodologic forms, product development, pragmatic problem solving *Range of stages*: ex ante, intermediate, ex post
2	Variability of criteria and indicators	**Two major approaches to quality assessment:** conventional metrics; indirect, field-based, and proxy criteria vs primary or epistemic measures of warranted interdisciplinary knowledge in the substance of the work **Expanded indicators:** for example, experimental rigor, aesthetic quality, new explanatory power, feedback to multiple fields, enhanced research capabilities, changing career trajectories, new public policies and treatment protocols, long-term impacts and unforeseen consequences
3	Leveraging of integration	**Key factors:** balance in weaving perspectives together into new whole, reaching effective synthesis, antecedent conditions for readiness **Criteria for leveraging and evaluating integration:** organizational, methodologic, and epistemologic components; strategies that promote communication and consensus; generative boundary objects
4	Interactions of social and cognitive factors in collaboration	**Requirements:** for example, calibrating separate standards, managing tensions among conflicting approaches, clarifying and negotiating differences among all stakeholders, compromising, communicating in ongoing and systematic fashion, engaging in mutual learning and joint activities
5	Management, leadership, and coaching	**Requirements:** managing tensions in balancing acts, consensus building, integration, interaction, common boundary objects, shared decision making, coaching the process **Categories of leadership tasks:** cognitive, structural, and processual
6	Iteration and feedback in a comprehensive and transparent system	**Requirements:** attuning a pluralism of values and interests, iterative work to insure collaborative inputs, transparency to include common stakeholding, feedback to the mission in a dynamic framework, mobility of participants, interaction and communication patterns
7	Effectiveness and impact	**Expanded indicators:** sensitivity to variety of goals in Principle 1 and variety of criteria and indicators in Principle 2; inclusion of unpredictable long-term impacts, returns on investment, value-added

experts is often smaller. In highly innovative work, developing validation criteria to gauge progress often becomes part of the actual process of inquiry.[7] The summary report[2] of the 2006 AAAS symposium cites a number of strategies in funding agencies, including creating "on-the-fly" electronic review teams, using "interpreters" who bridge the epistemic gap among content experts, asking candidates for grants to contribute the names of suitable peers, and forming joint panels and "matrix" schemes that combine disciplinary reviews with full-panel reviews among discipline-based and interdisciplinary members. Special funding programs may bypass conventional control mechanisms, but they run the risk of marginalizing interdisciplinary and transdisciplinary research.[2]

Lamont and colleagues' study[8] of fellowship competitions in social sciences and humanities furnishes a powerful analytical lens for thinking about interdisciplinary and transdisciplinary evaluation. Building on the work of Max Weber and Emile Durkheim, the team described the production of legitimacy that occurs in review panels. Review panels are "sites where new rules of fairness are redefined, reinvented and slowly recognized."[8] In the absence of customary rules, consensus on what constitutes a good proposal must be negotiated. Equilibria must be achieved between the familiarity and distance of non-expertise, between transparency and opacity, expertise and subjectivity, and between interdisciplinary appeal and disciplinary mastery. Methodologic pluralism is key to arriving at a judgment that is both consistent and limits bias.[8]

Finally, the logic of *measurement* returns the question of evaluation full circle to the gap between conventional metrics and the complexity of interdisciplinary and transdisciplinary research. Paralleling interdisciplinary studies and learning assessment, interdisciplinary and transdisciplinary research process and evaluation are grounded in the philosophy of constructivism. Appropriate evaluation is made, not given. It evolves through a dialogue of conventional and expanded indicators of quality. Traditional methodology and statistics have a role to play, but they are not sufficient. In the past, Sperber[10] admonishes, people seeking the legitimization of interdisciplinary initiatives had to be both parties and judges, educating their evaluators in the process of doing and presenting their work. The

emergent literature provides both parties and judges with an authoritative portfolio of methodologies, instruments, design models, guidelines, and conceptual frameworks anchored by a growing body of case studies and findings. They neither impose nor forestall evaluation awaiting a single-best or universal method that would be antithetical to the multidimensionality and context-specific nature of interdisciplinary and transdisciplinary work. They facilitate informed definition of the task and credible tracking of the actions and outcomes attendant to the substance, constitution, and value of the research.

References

1. Laudel G, Origgi G. Introduction to a special issue on the assessment of interdisciplinary research. Research Evaluation 2006;15:2–4.
2. Special issue on the assessment of interdisciplinary research. Research Evaluation 2006;15:1–80.
3. Feller I. Multiple actors, multiple settings, multiple criteria: issues in assessing interdisciplinary research. Research Evaluation 2006;15:5–15.
4. Klein JT. Afterword: the emergent literature on interdisciplinary and transdisciplinary research evaluation. Research Evaluation 2006;15: 75–80.
5. Langfeldt L. The policy challenges of peer review: managing bias, conflict of interests and interdisciplinary assessments. Research Evaluation 2006; 15:31–41.
6. Interdisciplinarity: problems of teaching and research in universities. Paris: Organization for Economic Cooperation and Development, 1972.
7. Boix-Mansilla V, Feller I, Gardner H. Quality assessment in interdisciplinary research and education. Research Evaluation 2006;15:69–74.
8. Lamont M, Mallard G, Guetzkow J. Beyond blind faith: overcoming the obstacles to interdisciplinary evaluation. Research Evaluation 2006;15: 43–55.
9. Laudel G. Conclave in the Tower of Babel: how peers review interdisciplinary research proposals. Research Evaluation 2006;15:57–68.
10. Sperber D. Why rethink interdisciplinarity? In: Interdisciplines: rethinking interdisciplinarity. 2003. www.interdisciplines.org/interdisciplinarity/papers/1.
11. Bruun H, Hukkinen J, Huutoniemi K, Klein JT. Promoting interdisciplinary research: the case of the Academy of Finland. Academy of Finland Publication Series 8/05. Helsinki: Academy of Finland, 2005.
12. Committee on Facilitating Interdisciplinary Research, National Academy of Sciences, National Academy of Engineering, Institute of Medicine. Facilitating interdisciplinary research. Washington DC: The National Academies Press, 2004. www.nap.edu/catalog.php?record_id=11153#toc.
13. Vickers J. [U]nframed in open, unmapped fields: teaching and the practice of interdisciplinarity. Arachné: An Interdisciplinary Journal of the Humanities 1997;4:11–42.
14. Stokols D, Hall KL, Taylor BK, Moser RP, Syme LS, eds. The science of team science: assessing the value of transdisciplinary research. Am J Prev Med 2008;35(2S):S77–S252.
15. Stokols D, Hall KL, Taylor BK, Moser RP. The science of team science: overview of the field and introduction to the supplement. Am J Prev Med 2008;35(2S):S77–S89.
16. Taylor BK, Moser RP, Trochim WM, et al. The evaluation of large initiatives (ELI) project at the National Cancer Institute: key findings and lessons learned. Proceedings of the NCI–NIH Conference on The Science of Team Science: Assessing the Value of Transdisciplinary Research. 2006 Oct 30–31; Bethesda MD.
17. Gray B. Enhancing transdisciplinary research through collaborative leadership. Am J Prev Med 2008;35(2S):S124–S132.
18. Stokols D, Fuqua J, Gress J, et al. Evaluating transdisciplinary science. Nicotine Tob Res 2003;5(1S):S21–S39.
19. Klein JT. Guiding questions for integration. Proceedings of the Integration Symposium 2004. Canberra: Land & Water Australia; 2005. CD-ROM #EC040735. land&waterAustralia lwa.gov.au.
20. Rosenfield PL. The potential of transdisciplinary research for sustaining and extending linkages between the health and social sciences. Soc Sci Med 1992;35:1343–57.
21. Aenis T, Nagel UJ. Impact indicator definition within a transdisciplinary research group. In: Tress B, Tress G, van dern Valk AJJ, Fry G, eds. Interdisciplinary and transdisciplinary landscape studies: potential and limitations. Delta Series 2. Wageningen, Netherlands: WUR Alterra, 2003.
22. Defila R, DiGiulio A. Evaluating transdisciplinary research. Panorama: Swiss National Science Foundation Newsletter 1999;1:4–27. www.ikaoe.unibe.ch/forschung/ip/Specialissue.Pano.1.99.pdf.
23. Krott M. Evaluation of transdisciplinary research. In: Encyclopedia of life support systems. Oxford: EOLSS Publishers, 2003. www.eolss.net.
24. Spaapen J, Dijstelbloem H, Wamlink F. Evaluating research in context: a method for assessment. 2nd ed. The Hague, Netherlands: Consultative Committee of Sector Councils for Research and Development, 2007.
25. td-net. Network for transdisciplinarity in sciences and humanities. www.transdisciplinarity.ch/.
26. Hirsch Hadorn G, Hoffmann-Riem H, Biber-Klemm S, et al. Handbook of transdisciplinary research. The Netherlands: Springer, 2008.
27. Bergmann M, Brohmann B, Hoffmann E, et al. Quality criteria for transdisciplinary research: a guideline for the formative evaluation of research projects. Frankfurt am Main: Institute for Social-Ecological Research (ISOE), 2005. ISOE-Studientexte, No. 13/English Version. downloads.isoe.de/literatur/evalunet_e.htm.
28. Boix-Mansilla V. Assessing expert interdisciplinary work at the frontier: an empirical exploration. Research Evaluation 2006;15:17–29.
29. Klein JT, Newell WH. Advancing interdisciplinary studies. In: Gaff J, Ratcliffe J, eds. Handbook of the undergraduate curriculum: a comprehensive guide to purposes, structures, practices, and change. San Francisco: Jossey-Bass, 1996.
30. Amey MJ, Brown DF. Breaking out of the box: interdisciplinary collaboration and faculty work. Greenwich CT: Information Age Publishing, 2004.
31. Derry SJ, Schunn CD, Gernsbacher MA, eds. Interdisciplinary collaboration: an emerging cognitive science. Mahwah NJ: Lawrence Erlbaum Associates, Inc., 2005.