GoodWork Project Report Series, Number 26

Assessing Interdisciplinary Work at the Frontier: An empirical exploration of "symptoms of quality"

by

Veronica Boix Mansilla and Howard Gardner

Address for Corresponsence

Veronica Boix Mansilla

Principal Investigator
Harvard Interdisciplinary Studies Project
Project Zero
124 Mt Auburn St.
5th Floor
Cambridge MA 02138

Phone: 617 496 6949
Fax: 617 496 9709

Paper presented at *Rethinking Interdisciplinarity*, A seminar supported by the C.N.R.S. project Society of Information, and moderated by Gloria Origgi and Christophe Heintz. December 1, 2003. We would like to thank The Atlantic Philanthropies for their generous funding of the Harvard Interdisciplinary Studies Project and the many researchers whom we interviewed for their illuminating insights.

I. Introduction

Arguably, the most dynamic research at disciplinary frontiers and in novel terrains is interdisciplinary.[i] Yet a re-emerging awareness of interdisciplinarity as a pervasive form of knowledge production is accompanied by an increasing unease about what is often viewed as "the dubious quality" of interdisciplinary work.[ii] Central to the controversy is the lingering challenge of assessing interdisciplinary work.[iii] Addressing the lack of available criteria to assess interdisciplinary work on its own terms, Donald Kennedy, Editor in Chief of Science, comments: "It is a terribly difficult issue….Interdisciplinary research institutes have the challenging task of producing as much good research as the departmental silos, judged however in a somewhat different [disciplinary] universe."[iv] "Criteria for judgment constitute the least understood aspect of interdisciplinarity," adds Julie T. Klein, "in part because the issue has been the least studied and in part because the multiplicity of tasks seems to militate against a single standard." [v] Against this background how does one determine what constitutes quality interdisciplinary work?

In this paper, we present the initial results of an empirical study of experts' views of interdisciplinary research. Specifically, we address the ways in which individuals in established and well regarded interdisciplinary research institutions assess the quality of their work and describe the dilemmas they confront. Our findings reveal that researchers systematically rely on indirect quality indicators (e.g., number of patents and publications-or type of journals and funding agencies associated to the work). Measures that directly address epistemic dimensions of interdisciplinary work (e.g., explanatory power, aesthetic appeal, comprehensiveness) proved rarer and less well articulated. In what follows, we introduce the study and summarize our findings. We delineate three core epistemic "symptoms" of quality interdisciplinary work emerging from our analysis: *consistency*, *balance*, and *effectiveness*.

II. Study overview

In the last two years, we conducted an exploratory study of research and teaching practices at exemplary interdisciplinary institutes and programs. Our goal was to understand qualities of expert interdisciplinary work in order to inform educational practice that fosters interdisciplinary understanding. In this study we defined "interdisciplinary work" as work that integrates knowledge and modes of thinking from two or more disciplines. Such work embraces the goal of advancing understanding (e.g., explain phenomena, craft solutions, raise new questions) in ways that would have not been possible through single disciplinary means.

In our formulation, disciplines comprise rich collections of theories, accounts, and findings believed to be acceptable within specifiable scholarly communities at a particular time. Such bodies of knowledge cannot be detached from the dynamic repertoire of methodological choices and forms of communication that give birth to them. We view interdisciplinary work as a purposeful means to reach a cognitive or practical goal (e.g., understanding, solving a problem) as opposed to as an end in itself. Our definition stipulates that disciplinary lenses be integrated in mutually informative networks of relationships rather than simply juxtaposed. By focusing on *disciplinary* integration –as opposed to the integration of multiple *perspectives*, disciplinary or not - our focus is more stringent than the "transdisciplinary" one presented earlier in this forum and in the literature.[vi]

The insights that we report stem from interviews with 60 researchers working in interdisciplinary institutes- specifically: the MIT Media Lab (ML), the Santa Fe Institute (SFI), the Center for the Integration of Medicine and Innovative Technologies (CIMIT), the Center for Bioethics at the University of Pennsylvania's (CB-UP), The Art-Science Laboratory (ASL), and the Research in Experimental Design group at XEROX-PARC (RED). We selected these institutions for their accumulated experience exploring novel disciplinary combinations (e.g., nonlinear dynamics and history; technology and music) and their good reputation. We expected that the difficulties associated with developing such novel integrations would have engendered a certain epistemological awareness among these researchers—a trait we were interested in capturing. Our

data consisted of in-depth, semi-structured interviews (including questions on how to assess interdisciplinary work), selected samples of work, and institutional documents.

III. Assessing interdisciplinary work – Challenges and measures

Most of the researchers in our study referred to the validation of interdisciplinary work as an obscure and challenging topic. They identified three sources of difficulty. First, they noted that disciplines themselves bring a variety of, often conflicting, standards of validation to the interdisciplinary meeting ground. Second, our subjects pointed to a lack of conceptual clarity about the nature of interdisciplinary work and its assessment, recognizing the need for a more systematic reflection in this regard. Third, they emphasized that in highly innovative work where novel territories are charted and few precedents are available, developing validation criteria is part of the inquiry process itself.

Faced with the task of making their assessment criteria explicit, researchers typically referred to *indirect* or *field-based measures*[vii] of quality. They pointed to indicators such as the *number* of accepted patents, publications, devices, and citations stemming from the work; the *prestige* of the universities, funding agencies, and journals in which it is placed; and the *approval* of peers and a broader community. "Simply counting things are easy answers as far as I'm concerned," claimed Jonathan Rosen, Director of the Office of Technology Implementation at CIMIT. "How many patents have you filed? How many patents have been licensed? How many new companies have been started? How many Science papers? How many Nature papers?" Field-based measures of this kind sidestepped the question of what constitutes warranted interdisciplinary knowledge by relying on social procedures of peer review, inter-subjective agreement, and ultimately consensus as generators of acceptable insight. Our subjects were often critical of these "proxy" criteria because they saw them as ultimately representing a disciplinary assessment of their interdisciplinary work. Yet they described these criteria as the standard way - however flawed - in which the quality of interdisciplinary work is determined at the forefront of knowledge production today.

When probed, most individuals also referred to more *primary* or *epistemic measures* of acceptability - i.e., epistemological indicators directly addressing the substance and constitution of the work. Researchers referred to a broad range of epistemological criteria (e.g., experimental rigor, aesthetic quality, fit between framework and data, power to address previously unsolved questions in a discipline). When considered collectively, these criteria shed light on three realms that demarcate *symptoms* of quality interdisciplinary work.

IV. Toward an epistemic framework for assessing ID work

Our interviewees highlighted the complexity of knowledge validation at disciplinary borders. In their view, interdisciplinary findings, theories, or exhibitions were not assessed as a sum of independent claims to be tested against equally independent disciplinary bars. Rather, researchers tended to provide a dynamic picture of knowledge validation in which the work as a whole can be assessed on three fundamental grounds:

1. the way in which the work stands vis à vis what researchers know and find tenable in the disciplines involved *(consistency with multiple separate disciplinary antecedents)*
2. the way in which the work stands together as a generative and coherent whole *(balance in weaving together perspectives)*
3. the way in which the integration advances the goals that researchers set for their pursuits and the methods they use *(effectiveness in advancing understanding)*

1. Consistency with multiple separate disciplinary antecedents

While the impetus of their interdisciplinary work was to move beyond established disciplinary boundaries, researchers often evaluated the degree to which their work was reasonably consistent with antecedent disciplinary knowledge (i.e., accepted methods, preferred conceptualizations, and epistemic values). They referred extensively to the act of satisfying multiple -sometimes conflicting- disciplinary standards at once.

In their view, borrowed disciplinary theories, methods, and communicative genres embodied epistemic values, which collectively informed the acceptability of interdisciplinary outcomes. For example, seeking to satisfy "two masters," SFI researcher John Padgett expected that his computer models of political life in Renaissance Florence would meet standards of scientific elegance and historical significance. Padgett valued work that could "explain coherently, highly heterogeneous phenomena... explaining heterogeneity with simple principles," while revealing important qualities of the period under study: "in a hundred years, will anyone read it? Historians care a lot about that."

The disciplinary canon was often a basic parameter against which researchers assessed their work. If a new finding was consistent with the "the laws of physics" or "current predictions in biology" it gained credibility. "There is a tremendous sense of freedom associated to breaking [disciplinary] rules" commented Mark Chow from RED, as he described his group's search for new ideas for an exhibit on "Experiments in the Future of Reading (EFR)" at the San Jose Technological Museum. At the same time, he added "you can do a lot of wild things, [but you need] somebody down the hall….who adheres to the scientific method [and is] squarely involved in the disciplines ….to say, well, this is against the laws of physics." If indeed interdisciplinary findings violated fundamental disciplinary tenets or revealed their limitations, additional justification was often seen as required. "The burden is on me to get a deeper understanding of their [disciplinary] methods and show them how their methods do and don't relate to our interdisciplinary methods" noticed Rosalind Picard, Media Lab Director of Affective Computing Research.

Ensuring appropriate fit between interdisciplinary products and findings and their antecedent disciplinary counterparts was not without challenges. Disciplines often conflicted vis à vis what they considered worth studying and what they viewed as warranted understanding. "What is this physicist doing writing a sociology proposal?" asked SFI researcher Mark Newman, as he imagined how colleagues in physics would critique his work on social networks. Illustrating differences in validation standards, SFI's Doyne Farmer commented: "Computer models are looked down on much more in economics than they are in physics. Mathematical proofs are

regarded as much more important in economics than they are in physics. Physicists are more comfortable with approximations."

Occasionally, standards stemming from different disciplines appeared as openly incompatible. For instance, the Experiments in the Future of Reading exhibit and the San Jose Technological Museum invited visitors to experience new forms of reading (e.g., interactive books, image and sound enhanced texts) while an explanatory voice guided them at each stop. XEROX-PAIR artist Paul DeMarinis spoke about what he and his colleagues perceived as the conflict between aesthetic and scientific dimensions of the exhibit. He claimed, "In art contexts you don't want a lot of text. You don't want to be told what it [the exhibit piece] is. You want it to come through and allow your mind to make other associations. In a science context you want to be sure that the person isn't misunderstanding what they are seeing." DeMarinis perceived the exhibit as embodying a central tension between science and postmodern theory in which "the [explanatory] text had the upper hand in formulating the theory [more] than anything you might experience yourself."

In sum, while a reasonable fit with antecedent knowledge in multiple disciplines strengthened the credibility of interdisciplinary outcomes, it clearly did not suffice as the sole source of rigor in deeming outcomes acceptable. Quality interdisciplinary understanding did not rest on a sum of established disciplinary rules, but rather on a unique coordination of disciplinary insights where disciplines played particular roles in the overall composition of the work. It is not surprising then, that our interviewees viewed reflective "balance" as a second symptom of quality interdisciplinary work.

2. Balance in weaving together perspectives

Assessing interdisciplinary work involves an appreciation of how disciplinary insights are intertwined and the relatively different roles that they play in yielding an overall composition. When disciplinary values conflict, compromises and negotiations are in order. Our interviewees valued work that exhibited a thoughtful balance of perspectives. Reflective balance did not imply an equal representation of disciplines in a piece of work, but a sensible one. For example, Arthur

Caplan, Director of the Center for Bioethics at the University of Pennsylvania, described the relative contribution of law and philosophy in his work. He illustrated how inquiry goals largely determined what counted as a workable balance.

"There is some tension between law and philosophy," he claimed, "as to what is the best way to talk, literally [about matters such as organ donation or human cloning]. Should we talk like lawyers and use case precedents and analogical reasoning? Do we use principles? That battle goes on. I think each [view] has a case and I think healthy tension is ok." Caplan crafted a pragmatic balance. "For certain issues you do want to know what really is the legal framework that you're operating in. And for some other issues like, 'Should we ban cloning?' — starting with the law is really not a good idea. For those, you really need to think philosophically about what cloning is and why it would be bad. You can make a law later." Caplan critiqued work that made legal recommendations "prematurely, before there is consensus about the values" as well as other cases where "there's a lot of consensus about the values and you don't need to dig in the same old ethical holes again".

Relatedly, our subjects referred to finding an appropriate balance vis à vis the levels of depth at which various disciplines were engaged. Again in this case, specific inquiry purposes seemed to inform the weighing of options against each other toward a sensible balance overall. In assessing the Experiments in the Future of Reading exhibit, Mark Chow noticed how the inclusion of an animal character in his piece made the content of the exhibit more accessible and interesting to his audience. He established that the exhibit succeeded "at being a dog that could read aloud, rather than a computer exhibit," adding that such success "allows you to relax some of the more stringent requirements of technology. For example, in terms of engineering, the performance of the reading was not 100% accurate, but it was, after all, only a dog."

Researchers found isolated disciplinary assessments of interdisciplinary work dissatisfactory because they failed to capture the knowledge composition as a whole— a critique often applied to peer review panels composed of specialists working in isolation. Donald Kennedy agreed; "Sometimes very good interdisciplinary papers may be viewed in a very negative light simply because narrow disciplinary criteria are used to assess them. If you have a paper that is

interdisciplinary and you think that it really does require a broad-gauged person, then I assume you try to find that kind of person. I think for reviewers it is hard not to make mistakes [when it comes to interdisciplinary pieces]."

As our interviewees described it, the interdisciplinary "balancing act" seemed to involve maintaining generative tensions and reaching legitimate compromises in the selection and combination of disciplinary insights and standards. Such a delicately balanced whole gained credibility if it did not violate central tenets of the disciplines involved. It gained relevance and acceptability if it afforded new understandings, solutions, products, and questions—including proposed transformations in established disciplinary practices. Determining the *effectiveness* of the leverage afforded by interdisciplinary integration was a most informative criterion to ascertain the success of interdisciplinary enterprises - the third symptom in our categorization.

3. Effectiveness in advancing understanding

Not surprisingly, researchers overwhelmingly tended to assess the success of their work in light of the aims of their inquiry. Interdisciplinary inquiries varied broadly in their specific aims and their favored validation criteria varied accordingly. When SFI physicists James Crutchfield and Mark Newman assessed their mathematical theories of innovation and network behavior respectively, they favored qualities such as their theories' ability to "*predict"* unstudied social and biological phenomena and their "tangible success in *explaining* something that wasn't explained by somebody else before." At CIMIT, the combination of physiology, molecular biology, nano-physics, and material sciences brought scientists like Joseph Vacanti closer to the creation of an unprecedented entity-- a vascularized artificial human liver that *"works*" and whose creation could have a "*transforming effect"* on organ transplantation surgical practice.

No single set of assessment criteria can do justice to the enormous variation in inquiry aims. Still it is worth noticing that, among our interviewees, contributions oriented toward pragmatic problem solving and product development seemed to place a premium on standards of "viability," "workability" and "impact". Contributions that seek formal algorithmic models of complex phenomena seemed associated to measures of "simplicity", "predictive power", and

"parsimony".[viii] Contributions aiming at a more grounded understanding of multidimensional phenomena (e.g. lactose intolerance or organ donation viewed in their intertwined biological, cultural, and psychological dimensions) tended to favor work that reached new levels of "comprehensiveness," "careful description," and "empirical grounding."[ix]

In addition to assessing the substantive leverage afforded by interdisciplinary work, researchers highlighted their methodological contributions. For example, Media Lab ethologist and artificial intelligence expert Bruce Blumberg claimed that his computer models of animal behavior provided novel method for cognitive scientists to "test out" their hypotheses. "Increasingly, computation is going to be a very valuable way to test out models [in psychology and cognitive science]," he claimed. "Because it is one thing to write a book and say this is how it [animal intelligence] must be organized. The proof is, could you take those ideas and actually implement them?" Enhanced methodological options, in turn, raised the standards for the interdisciplinary inquiry that followed. "What we used to do in the past are now things that 16 year-olds on the Internet can do," explained MIT's Rosalind Picard. "We were the only ones doing that 10 years ago."

Because of their non-paradigmatic approaches to knowledge production, our interviewees often confronted the challenge of a lack of precedents or viable contenders against which to compare their achievements. "We don't know if we are doing better than people working by themselves…we don't have these kinds of measures," noted Jonathan Rosen from CIMIT. Working in uncharted terrain implied that "there is no higher authority to appeal to adjudicate what's relevant knowledge and what's not or what is useful and not," claimed Anne Balsamo, Principal Scientist at RED, Xerox-PARC**.** "[When] you are at the cutting edge of anything, by definition you're taking risks that most do not take," added Joseph Vacanti. "So having somebody who can, in an expert way, help you is problematic because there's this vested interest problem, where the status quo and building on the status quo is most of what goes on." For these researchers, the effective advancement of interdisciplinary understanding involved not only developing new insights and methods, but also fashioning criteria with which to gauge their progress.

V. To Conclude

Under close scrutiny, researchers' views about the epistemic evaluation of their work revealed three realms in which to discern the acceptability of interdisciplinary work: (1) the degree to which new insights related to antecedent disciplinary knowledge, (2) the sensible balance reached in weaving perspectives together, and (3) the effectiveness with which a particular piece of work advances understanding and inquiry. As these researchers portray it, quality interdisciplinary understanding does not rest on an accumulated set of established disciplinary rules. Instead, each piece of interdisciplinary work revealed an idiosyncratic coordination of disciplinary insights geared to accomplish researchers' cognitive and practical goals.

Distilling workable criteria to assess the epistemic dimensions of interdisciplinary work requires that we tackle the problem at a productive level of analysis. Criteria too local (e.g., innovative experimental methods, accurate protocols, or rich original sources) will fail to account for the formidable diversity of aims and approaches legitimately characterized as "interdisciplinary." Categories too generic (e.g., coherence, accuracy, parsimony) will be ill-fit to capture the particular challenges associated with interdisciplinary integration. Categories with the greatest potential for the assessment of interdisciplinary work, our analysis suggests, will capture (1) the relationship between interdisciplinary outcomes and their multiple disciplinary antecedents, (2) the delicate adjustment that takes place as disciplines are intertwined toward a well-balanced whole, and (3) the leverage provided by the newly created hybrid insights.

In the end, while the assessment categories we propose might contribute to the cause of a more reasoned and reasonable consideration of interdisciplinary work, they will not render interdisciplinary work immune from "the unfortunate propensity for error" that characterizes human knowledge construction.[x] Indeed, interdisciplinary work gains its strength from its keen awareness of the provisional epistemic status of its findings. In our view, a serious assessment of interdisciplinary work should not seek to establish "warranted truths" nor, on the contrary, to let "all interdisciplinary flowers bloom." Such assessment should instead yield illuminating evidence to grant provisional credibility to the work in question. Thus the acceptance of an interdisciplinary insight (much like that of the framework here proposed) rests on the assumption

of the inherent provisionality of understanding and the endless human capacity to “retrench, retool, and try again”[xi]

VI. Notes

[i] Julie Thompson Klein, *Interdisciplinarity History, Theory, and Practice.* Detroit: Wayne State University, 1990. Diana Rhoten, "A Multi-method Analysis of the Social and Technical Conditions for Interdisciplinary Collaboration". In *Hybrid Vigor, Final Report. National Science Foundation*. September 2003. Nancy Sung et al "Educating Future Scientists" *Science 301* p. 1485. September, 2003.

[ii] Julie Thompson Klein. *Crossing Boundaries: Knowledge, Disciplinarities, and Interdisciplinarities*. Charlottesville VA: University Press of Virginia 1996. Michael Gibbons, et al *The New Production of Knowledge.* London: Sage 1994. Peter Weingart and Nico Stehr, Eds. *Practising Interdisciplinarity*. Canada: University of Toronto Press, 2000 .

[iii] Lisa R. Lattuca *Interdisciplinary Research and Teaching among College and University Faculty* Nashville: Vanderbilt University Press, 2001. Julie Thompson Klein, 1996 Ibid.

[iv] Interview excerpt November, 2002

[v] Julie Thompson Klein, 1996 Ibid, p. 210.

[vi] Helga Nowotny, "The Potential of Transdisciplinarity." In *Rethinking Interdisciplinarity* Christopher Heintz and Gloria Origgi moderators. Julie Thompson Klein et al, *Transdisciplinarity: Joint problem solving among science, technology, and society: an effective way for managing complexity.* Basel: Birkhauser Verlag, 2001.

[vii] Mihaly Csikszentmihalyi, "Society, Culture and Person: A systems view of creativity". In *The Nature of Creativity: Contemporary psychological perspectives* Robert Sternberg Ed. New York: Cambridge University Press, pp.325-339, 1988. In this chapter Csikszentmihalyi defines *field* as the social organization of a domain—i.e., as a network of interlocking roles (peers, reviewers, gatekeepers) in charge of selecting among proposed variations in a domain.

[viii] See examples of this kind of work in James P. Crutchfield, Peter Schuster Eds. *Evolutionary dynamics : exploring the interplay of selection, accident, neutrality, and function.* New York: Oxford University Press, 2003.

[ix] For examples, see William Durham. *Co-evolution: Genes, Culture and Human Diversity.* Stanford: Stanford University Press, 1991; and Rene Fox & Judith P. Swazey. *The Courage to Fail: A social view of organ transplants and dialysis.* Brunswick, NJ: Transaction Publishers, 2002.

[x] Catherine Z. Elgin. *Considered Judgement.* New Jersey: Princeton University Press, p. 12, 1996.

[xi] Catherine Z. Elgin, Ibid p. 12.