

Editorial

Integral Review (IR) continues to grow and evolve. For one thing we have begun to share the Editorial role among the Associate Editors, and Jonathan has handed to me (Tom Murray) the enjoyable task of welcoming you to our December 2009 issue. In this issue we are pleased to see, in addition to articles on a broad range of topics that I will describe below, a significant number of responses to Stein and Heikkinen's article in our June 2009 issue. I will present an extended overview of those responses and briefly discuss some of the threads they cover. First, I will begin with some news and announcements.

We would like to welcome two new members to our Editorial Board: Jennifer Gidley and Steve Wallis. Jennifer has previously published her work on evolving planetary consciousness with IR, and her extended article continues to be one of our most frequently downloaded. Steve Wallis has also published with us, and has also taken the initiative to edit, along with IR Editorial Board member Mark Edwards, a special issue of IR on *Developing Perspectives for Advancing Metatheory and Theory*. Prior to that, we will publish a special issue on *Toward Developing Politics and the Political*, edited by IR Associate Editors Sara Ross and Thomas Jordan. Later will be another, *Research Across Boundaries – Advances in Theory-building*, guest-edited by Markus Molz and Mark Edwards. If you want to propose a special issue for *Integral Review* you can inquire at ireditors@integral-review.org.

Many of us associated with *Integral Review* are looking toward the Integral Theory Conference next year at JFK University, and have submitted proposals. We also have plans in the works to host a gathering for friends of IR at or prior to the conference—we'll let you know through our *Occasional Newsletter* distributed to those who join IR's mailing list at http://integral-review.org/email_issue/index.asp.

In other news, IR has been picked up by EBSCO, the largest supplier of online journals to academic institutions around the world. IR is now listed in its humanities database, and we welcome the additional exposure this provides for a broader range of academic researchers to access works published in IR. We will be applying for listing in EBSCO's social sciences database in 2010, as well as the American Psychological Association's psycINFO database.

We have reviewed our submissions deadline policy in light of a number of factors and have now instituted a rolling submissions schedule. Thus submissions are accepted at any time, and are published in the issue following completion of review and any revisions. This reflects the reality that each submission requires individual treatment in terms of the time and effort it requires to reach publication, and it gives us more flexibility in organizing our editorial processes.

Looking to future issues and following from the themes explored in this issue, we are particularly inviting papers that further explore the themes of assessment of adult development and the application of integrally-informed models for organizational and leadership change and transformation.



In This Issue...

As usual, this issue features articles on a wide range of leading edge interdisciplinary topics covering the theory-to-practice spectrum. Several submissions are in response to Stein & Heikkinen's paper on *Models, Metrics, and Measurement in Developmental Psychology* in the June 2009 issue of Integral Review. I will report on these in the succeeding section.

The section of peer reviewed articles in this issue begins with *The Coherent Heart*, in which Rollin McCraty and colleagues share a detailed report of a wide range of fascinating research studies during the last decade that show how emotion and the physiological state of the heart interacts with numerous physical and psychological systems. They show how coherence in heart rhythm and harmonious synchronization in heart-brain interactions are associated with positive emotions and significantly impact a number of indicators of well-being. They work with the Institute of HeartMath, which has developed contemplative activities for the practice of sustained positive emotion that induce heart rhythms related to health and well-being. The research uses dynamic systems theory and looks at both neurocardiological data and the heart's generation of electromagnetic fields.

In *The Ethics of Promoting and Assigning Adult Developmental Exercises: A Critical Analysis of the Immunity to Change Process*, Sofia Kjellström shares her experiences of using Kegan and Lahey's Immunity to Change Process in higher education. This process is usually used in organizational contexts to help participants uncover unexamined or unconscious assumptions that prevent them from reaching important goals and commitments. It requires participants to reflect in deep and vulnerable ways upon disparities between intention and action. Kjellström discusses a number of limitations and ethical issues that arise in trying to use the method in educational contexts and in so doing introduces her perception of hidden assumptions of the Immunity to Change model itself. She examines the possible mismatch between organizational contexts and classroom contexts in terms of the desire to change and reflect upon oneself, commitment to the change process, trust among those engaged in the process, and the developmental level required to use the process. The results are of interest to all who facilitate the process, and those who value giving students authentic, engaged, and/or transformative experiences in the classroom, but must grapple with the fact that school has critical differences with real life contexts such as work and family.

Our third peer reviewed article is Bill Torbert and Reut Livne-Tarandach's *Reliability and Validity Tests of the Harthill Leadership Development Profile in the Context of Developmental Action Inquiry Theory, Practice and Method*. This is a response to Stein & Heikkinen's paper, and is summarized with the other responses later.

The editorially reviewed section of this issue begins with *Toward Integral Higher Education Study Programs in the European Higher Education Area: A Programmatic and Strategic View*, in which Markus Molz reviews the state of higher education in the EU, looking at trends and imagining possible futures. He sets the historical, pedagogical, and institutional context for attempts at bringing integral approaches to higher education, and argues for the critical importance of such approaches, despite the only moderate and often ungainly attempts to date.

His analysis of why it has been so difficult to create sustained integral programs leads into a series of recommendations for future attempts.

In *Consciousness in Evolution Sketch for a New Model – A Speculation* Donald Padelford takes us on an informal yet in-depth tour of his "Twenty Tenets of an New Model of Evolution." Ken Wilber proposed an earlier, different set of Twenty Tenets in *Sex, Ecology, Spirituality*, but Padelford introduces his own insights and includes references to more recent work in the field. He brings a unique voice to the oft-pondered relationships between mind, matter, hierarchy, development, consciousness, causality, and interiority. His analysis is based in part on the construct of hierarchical negentropic systems.

R. Scott Pochron's *A Leadership Journey: Personal Reflections from the School of Hard Knocks* is a reflection upon the lessons learned in many years of leadership in diverse organizations. His themes include "organization as organism," "leading as enacting an emerging future," "disequilibrium," and "leadership as a state of mind." Along the way, he draws stories and principles from leadership and organization change authors including Quinn, Jaques, and Senge.

Russ Volckmann graces our pages with another engrossing interview. In this issue he speaks with Alfonso Montuori, Chair and Professor of the Transformative Studies Ph.D. and Transformative Leadership Master's program at the California Institute of Integral Studies, and IR Editorial Board member, on the themes of creativity and transdisciplinarity. These themes emerge and resurface as Montuori discusses his stint as a professional saxophone player in England, his involvement with the Monterey Institute of International Studies, his teaching in China, and his work editing the *World Futures Journal*, taking us on an international tour of the places he has lived and worked. It is a paradigmatic yet idiosyncratic ride through the cultural shifts in this age of Aquarius, and a conceptual voyage through theories of everything, general theories of leadership, metatheory, epistemology and complexity, disciplinary promiscuity, and "liquid life." In his interview Volckmann helps reveal that life of "creativity...as a way of being" with each movement echoing "the sound of surprise."

Editorial Board member Roland Benedikter reviews John Holman's book, *The Return of the Perennial Philosophy: The Supreme Vision of Western Esotericism*. While he appreciates its treatment of trends in exploratory first person "empirical spirituality" vs. traditional faith-based spirituality, overall he finds the book to be "an interesting but incomplete contribution," in part due to the author's inability to take enough of an objective and non-promotional distance from his subject.

Responses to Models, Metrics, and Measurement in Developmental Psychology

In our June 2009 issue we published an article by Zak Stein and Katie Heikkinen titled *Models, Metrics, and Measurement in Developmental Psychology*. When we first spoke with Zak about publishing a paper on this topic we were aware that it could serve as a lightning rod for vigorous debate, and saw this as an opportunity to plan for our first article-response-rejoinder format crossing multiple issues of IR. The obfuscated lines between research and the

marketplace, problematic in so many academic fields, show up vividly in developmental psychology, where developmental assessments are being used increasingly to evaluate development of individuals in domains including leadership/management, therapy, and academic programs.

Through various channels, we invited responses from theorists and practitioners, and are pleased to publish in this issue a number of responses, including submissions from Bill Torbert, Michael Commons, Theo Dawson, Michael Basseches, the principals at Ikan, myself, and separate rejoinder articles from Stein and Heikkinen. Stein and Heikkinen's earlier article had also inspired Bill Torbert to host a face to face meeting of developmental practitioners and theorists. Terri O'Fallon reported her reflections on that meeting in the October issue of *Integral Leadership Review*, (linked in the Table of Contents for this section), and we publish a transcript of that meeting's closing conversation, which gives a flavor of the day-long meeting.

Summary of Key Elements of the Original Article

In *Models, Metrics, and Measurement*, Stein and Heikkinen introduce a framework and terminology aimed at supporting more rigor in developmental models and metrics (Figure 1 in the paper). They differentiate "soft metrics" from "calibrated metrics" with the latter having more stringent "reliability and validity profiles." They also propose that every metric we use for developmental assessment or research should have reliability and validity profiles associated with it. "If we want to see an integral and developmental worldview gain a real institutional foothold—radically reforming business, government, education, therapy, and our own sense of human potentials—we need to get serious about our quality control standards" (p. 19).¹

Central to Stein and Heikkinen's 2009 paper is a strong proposition that only *calibrated* metrics are appropriate for assessing individual performance, as only they can make "reliable fine-grained distinctions." They propose that soft metrics are admissible for research studies, whose results are averaged over many subjects, and where the outcomes do not reflect upon any individual. When soft metrics used in research are adapted to evaluate individuals, the margins of error are large (according to the very definition of soft metric). For such metrics, an assessment outcome of, for example, "you are at level 4" is misleading and potentially harmful, when a more accurate report might be "an 85% certainty that you are within levels 3, 4, or 5" (my example, not a quote from the paper).

Stein and Heikkinen go on to do a "limited and exploratory literature review of metrics currently in play," looking for published evidence of validity and reliability quality assessments. The paper's Table 2 summarizes this review, and includes work by Torbert, Cook-Greuter, Jaques, Kegan, Beck & Cowan, Graves, Commons, and Dawson (and their colleagues). The results "reveal a conspicuous lack of psychometric rigor on the part of some of the most popular developmental approaches" (p. 4). The only metrics that rise to the level of calibrated

¹ This article continues to take up issues of praxis ethics, following papers at 2008 Integral Theory Conference and in *Integral Leadership Review*, in which Stein and Heikkinen suggest that there is a persistent fact/value confusion in the field of developmental research and application, in which higher levels of development are assumed, implicitly or without sufficient justification, to be desired and of superior value.

measurement are those used in the work of Dawson and colleagues (which draws on Commons' and Fischer's developmental models). "It appears that the LAS and HCSS are the only metrics that have been calibrated using quantitative indexes of internal consistency" (p. 19). Stein is transparent about the fact that he is Dawson's partner in a commercial venture based upon Dawson's LAS scoring system.

Pointing to the non-exhaustive nature of the literature review (they searched in Google Scholar and PsycINFO using a reasonable set of keywords), the authors say they are "just trying to kick-start a concerted discourse about quality control in the field" and they invited colleagues to "work together to fill out and expand the [publications list], making clear to each other and to all those affected exactly what we are doing" (p. 19).

The Invitation

At *Integral Review*, we are interested in fostering theory/praxis integration and see acts of publication themselves as opportunities to investigate integral principles. Stein and Heikkinen's prescriptions are far reaching and their analysis is provocative. They call upon individual scholars and the field as a whole to take a hard look at the ethical implications of their work. As the lines between academic research and commerce increasingly blur in all scholarly fields, such questions become both more critical and more harrowing to address with authenticity and openness. In our invitation for responses we said we were interested in submissions that

integrate first-, second-, and third-person perspectives, and illustrate an author's embodiment of higher (second tier, integral, metasystematic, etc.) levels of ethical and psycho-socio-systemic reflection. We believe that a community of "integral" scholars (however we might define it) has the potential to engage in dialogs that embrace authenticity/vulnerability/empathy *and* rigor at levels rarely seen in academic discourse.

In other words, to what extent can the scholarly author or the business practitioner (so often one and the same these days) openly reflect upon and be transparent about deeper motives, possible biases and blind spots, emotional overtones, and uncertainties? Can we as colleagues (and often as friends) in dialog pose insightful yet caring questions that support integrity and transparency without falling into patterns of critique that feed more destructive than constructive to the recipient? This challenge is romantic and naïve if not tempered by an acknowledgement of the acute difficulties and complexities it raises. But we have hopes that articles in the integral community might illustrate some small but significant movement at this leading edge. The in-person dialogs at Torbert's meeting certainly seemed to illustrate this potential. We will leave it to the reader to judge how well the respondents did so.

Responses to the Article

We were pleased to accept a number of responses, introduced below. Our first response came in an informal letter from Bill Torbert which was published in IR's Occasional Newsletter #7, July 2009, along with a reply from Stein (see link in the Table of Contents). We also want to note Terri O'Fallon's piece in *Integral Leadership Review* (also linked in the Table of Contents)

which is a description of the October 3rd meeting at Bill Torbert's house, which arose as a response to Stein and Heikkinen's article.

We begin this section of responses with a transcript of the *Closing Conversation of the "Validity Day" Meeting at Bill Torbert's, October 3, 2009*. This transcript portrays a rich cast of characters deeply engaged in an intense and generative conversation that was the culmination of this meeting. It addresses the challenges ahead as the field of developmental psychology enters into a time when such psychometrics are becoming more in demand and when practitioners may not have the depth of rigorous knowledge felt necessary to use such tools appropriately. Thus an intention emerged for a knowledgeable group to form a community of inquiry to better lay some critical foundations for how this future unfolds.

Bill Torbert and Reut Livne-Tarandach sent an in-depth (19 page) reply (noted above in our peer review section) in *Reliability and Validity Tests of the Harthill Leadership Development Profile in the Context of Developmental Action Inquiry Theory, Practice and Method*. Torbert and Livne-Tarandach "challenge [Stein & Heikkinen's] critique of the LDP as a soft measure unsupported by published, quantitative psychometric reliability and validity studies ... and present both previously unpublished and previously published-but-not-aggregated studies illustrating Harthill LDP as a well-calibrated measure of adult ego-development." They also "consider what a social science and a social practice based on the developmentally late action-logics will look like [once it embraces] 1st-, 2nd-, and 3rd-person research and action with co-participants in live settings."

Theo Dawson sent two brief responses (3 pages) that we combined into *Working Within the Limits: Thoughts on Stein and Heikkinen*. She reflects upon the nature of assessment validity and reliability and upon the constraints upon real-life use of such instruments. She included an appendix titled *Reliability – Some Basics* to help our readers understand some of the subtleties of the topic. She comments on the practical use, or lack thereof, that developmental measurements may have for end users, noting that "the more we hone in on a measureable dimension—the greater our precision becomes—the narrower the construct becomes." She continues by saying that "Formal measures, no matter how many great ones we design, should always be employed by knowledgeable mentors, clinicians, teachers, or coaches as a single item of information about learners that may or may not provide useful insights into their needs."

In *A Personal Counterpoint to Stein and Heikkinen* (3 pages) Michael Basseches suggests that in addition to questions of validity and reliability, it is equally important to question the use to which developmental assessments are put. He takes Dawson's concerns further and cautions against the potential of a "tyranny of measures that replaces respectful discourse and collective adaptation as the social context." He uses the developmental capacity of dialectical thinking from his own research to illustrate. Only when one has "encountered the limits of [their] lower level structure [i.e., formal operational thought], and begun to glimpse the power of the novel alternative to transcend those limits [i.e., dialectical intelligence]" is one in a position to appreciate the implications of an assessment of dialectical intelligence (or any higher stage capacity).

In *Further Issues in Stage Metrics* (5 page response, plus 5 pages of References and Appendices) Michael Lamport Commons takes the opportunity to clarify the nature and history of the Model of Hierarchical Complexity—the construct and the surrounding theory. Zak's own work with the LAS (with Dawson) draws strongly from Commons' work, so this response provides a deepening background context for Stein and Heikkinen's paper. Commons also differentiates his work from Fischer's Skill Theory, which also contributed substantially the LAS.

From the commercial side of the street comes a response from the principals at InterKannections (ikan.biz). *A Practitioners' Perspective on Developmental Models, Metrics and Community* (19 pages) by Chad Stewart, Zach Smith and Norio Suzuki, gives a peek at a range of assessment instruments being used by a particular consulting company. The authors describe integrally-informed instruments that they have developed or adapted for assessing a range of constructs important to businesses, including 360 performance evaluations, leadership profiling tools, and an organizational sustainability scorecard. They conclude that purely linguistic evaluation instruments must be supplemented by performance-based assessments and metrics that evaluate outcomes (lagging indicators) as well as internal capacities (leading indicators).

Through the ensuing debate I was moved to pen a response titled *Intuiting the Cognitive Line in Developmental Assessment: Do Heart and Ego Develop According to Hierarchical Integration?* in which I explore some of the underlying issues in a question that came in these debates regarding developmental assessments: are they too tied to "the cognitive line" (or language skills)? In this informal analysis I inquire into whether the development of capacities related to heart, ego, and spirit can be captured using current models based on hierarchical integration and coordination.

In his rejoinder *Educational Crises and the Scramble for Usable Knowledge*, Zak Stein carries forward his appeal for more serious discussion among the designers and users of developmental models and tools, this being a third installment in what he intends as a series of papers on the theme. He argues that the profound challenges facing global society today "have critical educational dimensions" yet the "pressing global demand for innovative educational solutions... has the potential to systematically distort the production of relevant usable knowledge" about human learning and development. He iterates his previously-expressed support for Sara Ross's 2008 call for an "Institute for Applied Developmental Theory" focusing on quality control and fostering collaboration." He goes on to identify issues within the community of developmentalists that he sees as inhibiting the quality of dialog necessary to make sufficient progress.

In her rejoinder *It's an Empirical Question: On Cognition and Ego*, Katie Heikkinen jumps bravely into rarely explored territory in warning against the phenomena of disqualifying the arguments of a colleague because of perceived developmental inadequacies (coming from a lower developmental position). She follows Stein's rejoinder comments section on *Lack of Epistemic Humility*. This "labeling colleagues" territory is rich, sensitive, and nuanced. In part due to some of Wilber's arguments and others mentioning "adequatio," it has become common doctrine in the integral community that one can't "see" stuff at higher levels than they are at, at least along a particular developmental line. And differences in expertise in any domain can be (is

not always) hierarchical in nature. But putting this intuition to use with real people in real contexts, and fleshing out which elements of it are true, useful, and/or ethical, is problematic. In addition to raising questions of how we could, or whether we would ever want to, "objectively" assess scholarly adequacy in a domain, it raises the question of whether we can profess and value humility, yet be able to compassionately and squarely look at deeply ingrained forms of "arrogance" that are tacitly cultivated in academic work. In a sense we know "labeling colleagues" is wrong, and Heikkinen and Stein help articulate why, but when we catch ourselves engaged in this type of thought, individually (I certainly have) or systemically, what options are available to us in open dialog and self/system analysis?

Heikkinen's discussion on this theme is limited to a particular issue, but she and Stein open a Pandora's box that may hold extensive treasures should the community take up this delicate issue in more depth and breadth. The issue points to a possible shadow element of the integral community of theory and practice. As was said in IR's invitation for responses "such questions become both more critical and more harrowing to address with authenticity and openness," yet we believe that a community of "integral" scholars (however we might define it) has the potential to engage in dialogs that embrace authenticity/vulnerability/empathy and rigor at levels rarely seen in academic discourse." Any takers? Write to us!

Anyways, back to Heikkinen's rejoinder. Following her discussion on "labeling colleagues" she asks some frank questions about whether any of the respondents addressed the core questions about calibrated metrics. She goes on to explore questions raised in my own article on whether wisdom skills can be evaluated using hierarchical complexity. Finally, she describes her current research project involving recoding Kegan's Subject-Object Interview data using the Hierarchical Complexity and the LAS.

Concluding Thoughts

In the end we see that Stein and Heikkinen's June 2009 paper served well to motivate a wide range of responses on the general theme the appropriate use of developmental assessments, and to further the dialog on this important topic. Their call for assistance in filling in their table of published findings on the validity of developmental instruments was answered only by Torbert. Issues surrounding the practical usefulness and valid application of developmental tests were widely addressed, but Stein and Heikkinen's core concern about whether we are doing more harm than good in offering clients measurement results that are not transparent about their degree of uncertainty was not picked up by many.

Test validity and reliability are very time-consuming to establish, and require access to controlled situations and large numbers of data points. They also require instruments that remain fixed (that the measurement tool is constant), which prohibits the type of ongoing-improvement that characterize most practical methodologies. Many instruments capture a particular model for categorizing human capacities (such as the Myers-Briggs assessment) that have proven to be "tried and true" useful in the field for "adding value" to clients and customers, but have not been proven to have the psychometric properties the Stein and Heikkinen claim are necessary for accurate (thus ethical) use. Is using such instruments like the story of the man who lost his keys

in the bushes but was looking for them under the street lamp because "the light was better there"?

But it is not that clear cut. One thing that several of the authors working in commercial contexts allude to indirectly is the potential for these instruments to serve as meaning-generating and self-reflection tools. Given a categorizing scheme, whether it be Action-logics, Enneagram types, or Astrological types, one can assess individuals (or groups) and present the scheme as a conversational tool for framing weaknesses, strengths, goals, opportunities, and intentions. If the uncertainty of any measurement is clarified, perhaps supplemented by self-evaluation ("what category do you think you are in?"), then instruments that lack strong validity or reliability can nonetheless be quite useful. On the other hand, repeated client feedback that the consultation seemed insightful and quite useful can mislead the practitioner into thinking that the tool is valid and reliable, and no substitute for rigorous psychometric evaluation.

The responses suggest that we may benefit from differentiating two contexts. In cases where assessments are used by third parties to make judgments that will have an effect on the lives of individuals, great care must be taken to avoid a "tyranny of measures." On the other hand, when the instruments are used by individuals or groups in dialogic, self-reflective first-, second-, third-person meaning-generation processes, the validity and reliability standards can be relaxed and still prove ethical and useful.

Responses to Stein and Heikkinen's original paper, including Stein and Heikkinen's Rejoinders, may have raised more questions than they answered, and to me it feels as though on many topics we are still in mid or early dialog, as a community of inquiry. We invite further thoughts, responses, and articles from our readers, which we may publish in our next newsletter, or our next Issue.

I hope you enjoy this issue!

Tom Murray
Associate Editor
December 12, 2009