

Virtuous cycles of learning: using formative, embedded, and diagnostic developmental assessments in a large-scale leadership program

Zachary Stein, Theo Dawson, Zachary Van Rossum

Lectica, Inc.

Sandra Hill, Joel Rothaizer

Clear Impact Consulting Group, Inc.

Introduction

In this paper we present preliminary results from a series of ongoing action-research projects involving the use of embedded diagnostic developmental assessments (called LectaTests) in leadership education contexts. These findings are presented to support a particular meta-theoretical approach to learning and education in which embedded assessments form a crucial part of ongoing *virtuous cycles* of action, feedback, support, and learning. We present two types of evidence. First, we compare developmental growth across 8 program evaluations in which LectaTests were and were not embedded. Second, we examine how embedding LectaTests in a large-scale leadership development program affected the growth of managers and their direct reports. We review these findings with an eye toward detecting the benefits of using developmental assessment as embedded diagnostics alongside their use as research instruments. We begin by using Integral Theory to structure a discussion about the ideal function of developmental assessments in educational contexts.

Virtuous cycles of learning: educational meta-theory and developmental assessment

In this paper we argue that developmental assessments provide the greatest benefit when they are used to support learning. When they are designed according to certain meta-theoretical principles, developmental assessments can be used both to objectively and accurately measure outcomes (i.e., as summative assessments) and to provide diagnostic feedback and educational support (i.e., as formative assessments). Designing developmental assessments to ensure objectivity and reliability increases their value for research and program evaluation, while designing them to serve as richly diagnostic educational aids transforms the space of pedagogical possibility and re-frames the meaning of testing from coercive to persuasive. LectaTests combine the formative and summative functions of assessment, which allows them

to facilitate multi-level learning processes involving learners, educators, administrators, researchers, and assessment developers.

The idea of using developmental assessments to facilitate multi-level learning brings the approach discussed here in line with a variety of meta-theories and philosophies. Integral Theory (Wilber, 1999) and Dialectical Critical Realism (Bhaskar, 1991), the foci of this conference, are two of the most directly relevant. These two philosophical systems share the idea, along with other great systems, like those of Habermas (1984) or Peirce (1865), that *learning is an intrinsic feature of social life*. They stand opposed to the dominant ideology of learning that guides most human capital management and education, where learning is a kind of *imposition* on social life and students are worked on from the top down, having knowledge put into them, only to be tested to see how much they have retained. Instead, the philosophies of learning that inspire the work presented here suggest that educational processes ought to be understood as dialectical processes, involving iterative cycles of action, feedback, and reflection. This kind of *virtuous cycle of learning* is a universal pattern immanent in both biological life and human social interaction (Piaget, 1971).

This is the central idea in the meta-theory of education that is the focus of our discussion: learning takes place through *virtuous cycles* of reflective agency and objective feedback. From an infant's first call and response with its mother to the scientist's iteratively revised hypotheses, learning involves virtuous *action-feedback-reflection* cycles, which result in the creation of increasingly complex and adequate sets of skills and concepts. This idea, which Baldwin (1906) first investigated under the heading of the "circular reaction" has roots in the earliest biologically oriented psychologists, who documented its trans-species validity. Piaget (1952) brought the notion into developmental psychology in the 20th Century. Early dynamic

systems theorists (Waddington, 1966) applied the notion to organisms and the post-war cybernetic revolution to organizations (Forrester, 1964). Below we describe a particular Neo-Piagetian interpretation of this perennial idea, one that situates learning in dynamic socio-cultural contexts and motivates the use of developmental assessment technologies to catalyze virtuous cycles of learning across multiple levels in educational organizations.

Virtuous cycles by design: the model and metric behind LectaTests

The approach discussed here is based on the work Kurt Fischer and Theo Dawson, who have worked to preserve the Baldwinian-Piagetian insight that human development is a constructive and dynamic process that unfolds across a series of hierarchically nested levels. Fischer's expanded upon this view, describing a developmental sequence that can be observed across contexts, domains, and time scales. His model has been employed to describe the variability of development and to build domain specific learning sequences (or developmental pathways). The developmental levels of Fischer's Dynamic Skill Theory (1980) are understood as major re-organizations of behavior that mark the use of qualitatively new kinds of capabilities—higher-order capabilities built upon combinations of previously constructed lower-order capabilities. These levels unfold over the course of the entire lifespan, ranging from *reflexes* and *sensorimotor action-schemes* through *concrete representations* and *abstractions* to overarching *principles*. The construction of new types of capabilities—moving from one level to the next—involves the active coordination and hierarchical integration of existing skills. This process requires effort, resources, and social relationships and is thus greatly impacted by contextual factors, such as education (Dawson-Tunik, Commons, Wilson, & Fischer, 2005; Kurt W. Fischer & Bidell, 2006). The levels of Dawson's (Dawson, 2010) metric, the Lectical[®] Assessment System (LAS) (Dawson, 2010) are equivalent to Fischer's skill levels

In the 1980s, domain-specific developmental assessment systems proliferated. Kohlberg's (1969) *Stage and Sequence* had ushered in decades of longitudinal research designed to reveal the sequences through which skills and concepts develop over the course of the lifespan within particular knowledge domains. Debates about domain specificity became entrenched and each time a new domain of knowledge was studied a new assessment system was devised. In this context, Dawson, drawing on the domain-general models of Fischer and Commons, began a series of cross-metric comparison studies aimed at isolating the latent developmental dimension underlying various domain-specific developmental assessment systems (Dawson, 2002; Dawson-Tunik, 2006; Dawson-Tunik et al., 2005). The result was the specification of domain general, deep structural aspects of performances and tasks that can be taken as indices of development. That is, Dawson's cross-metric comparison studies suggest that many of the different domain-specific assessment systems—particularly those bootstrapped from longitudinal data sets—share a common developmental dimension.

This underlying developmental dimension—best characterized in terms of certain deep structural aspects of performances—is represented in the levels identified by the LAS. These levels, which as noted above, are equivalent to Fischer's skill levels (described in more detail below). Importantly, because the LAS targets domain-general deep structural aspects of performances, it allows for a rigorous separation of the structure of a performance—which is indicative of its developmental level—from the content of a performance—which often reliably co-varies with level but is not an index thereof. This clear differentiation of structure from content is a critical aspect of the overall method, and provides numerous methodological and theoretical advantages over domain-specific scoring systems (Dawson-Tunik, 2004). The advantage most relevant in this context is the ability to identify the full range of within-level variability in content. For example, in the domain of leadership decision-making, some

perspective-taking and seeking behaviors once thought to be developmental in nature, turn out to be independent of developmental level (Dawson & Stein, 2011). It is impossible to detect effects of this kind when aspects of perspective taking and seeking are viewed as indices of development, as they are in some developmental assessment systems (Cook-Greuter, 1999; Kegan, 1982; Loevinger, 1966). Because the LAS does not employ content-based scoring criteria, it can be used to ask questions about the relation between developmental level and specific skills or conceptual content in any knowledge domain. As explained below, the LAS is the common core metric behind all LectaTests.

LectaTests

For purposes of the discussion here, the Letical Decision Making Assessments (LDMA) will serve as an example of what LectaTests are and how they work. The LDMA is an online assessment of how people make decisions in leadership and management contexts. It is designed for students, leaders, and managers, and individuals who are thinking about moving into leadership and management, and was originally developed as part of a project investigating the development of leadership skills in a U. S. federal government agency (Dawson & Stein, 2004).

The LDMA focuses on 8 distinct aspects of decision-making, including collaborative capacity, contextual thinking, and cognitive complexity. It presents a common workplace dilemma that involves conflicting interests, then asks the test-taker—through a series of standard probes—to discuss the nature of the problem, describe two possible solutions, compare these solutions, and describe an ideal decision making process for similar situations.

There are many LDMA dilemmas, all of which are designed around real-life management situations that involve the coordination of hierarchically nested perspectives, including those of

a protagonist and his or her subordinates (as individuals and teams), a supervisor, senior management, an institution, and people served by the institution. Test-takers are either assigned a dilemma or are asked choose their own dilemma from a drop-down list. They are then asked to unpack their thought process by discussing the factors they would consider in addressing the dilemma, the actions they would take to solve the problem, and their decision-making process. One of the most commonly used dilemmas is the *Office Reorganization* dilemma:

“You have been a manager in one of the most technically savvy and productive offices in the company for the last three years. Almost 80% of the employees have at least Masters degrees and many have doctoral degrees in engineering or computer science. This has been much easier than your last management position, because here you have such great respect for the ability and drive of your employees. When your supervisor retired 3 months ago, the senior leadership team decided to replace her with an executive hired from outside the company. The individual that was finally selected after a lengthy interview process has only been on the job for 1 week and is already stirring things up. After his first walk-through of the spaces, essentially a large cubicle farm, he announced that he was going to redesign the space to "open things up" and encourage greater collaboration and exchange of ideas among members of the group. You have been presented with a drawing of how the space will be reconfigured and a very aggressive time-line for the work, both of which you share with your employees. This normally quiet, reserved group is visibly outraged. How can they be expected to do highly technical work without the quiet and privacy of their cubicles? What's wrong with using a conference room when collaboration is called for? They are looking to you to stand up for them.”

This dilemma, like all the LDMA dilemmas, is purposefully open-ended and *ill-structured*. In other words, like the classic Kohlbergian dilemmas, these dilemmas are amenable to multiple interpretations and thus can be understood and approached from multiple developmental levels. So there is no “right” answer, only more or less developed responses displaying greater or lesser degrees of complexity, sophistication, coherence, perspective taking, and other aspects of reasoning and content. And as we show below, even respondents at the same developmental level do not see the same things as relevant or take the same variables as salient. So, the dilemmas are sufficiently complex to allow for a wide range of variability in performance. The evidence to date, based on over 1500 LDMA’s, reveals no differences in the difficulty of different dilemmas with respect to the Lectical Scale™ (Dawson & Stein, 2004, 2006).

The LDMA, like all LectaTests, is scored with the LAS, which provides explicit criteria for determining the complexity level and phase (degree of elaboration within a given level) of performances or tasks in any domain of knowledge (Dawson, 2010). Its most commonly used levels consist of the last eight of thirteen complexity levels (single representations, 6; representational mappings, 7; representational systems, 8; single abstractions, 9; abstract mappings, 10; abstract systems, 11; and single principles, 12) corresponding definitionally to Fischer’s (1980; Kurt W. Fischer & Dawson-Tunik, 2006) skill levels. Its phases (transitional, a; unelaborated, b; elaborated, c; and highly elaborated, d) are based on empirical evidence regarding the way learning within levels progresses. This evidence has been derived from a large database of scored interviews and essays (Dawson & Wilson, 2004). Scores are represented with the level first and phase following. For example, elaborated abstract mappings is 10c.

LAS scoring procedures are partially derived from Commons' (Commons et al., 1995) General Stage Scoring System (GSSS). Like the GSSS, the LAS is designed to make it possible to assess the complexity level of a performance based on its level of differentiation and integration—deep structure—without reference to its particular conceptual content. Rather than making the claim that a person occupies a level because he or she has, for example, elaborated a particular form of perspective taking, the LAS permits us to identify performances of a given complexity level and then to ask (empirically) what the range of perspective-taking forms are at that complexity level. Thus, it avoids much of the circularity of many stage scoring systems (Brainerd, 1993), such as the Perry (1970) scheme, Colby and Kohlberg's Standard Issue Scoring System (Colby & Kohlberg, 1987), Kegan's Self-Object Interview and Scoring System (Kegan & Lahey, 2002), and the Reflective Judgment Scoring System (King & Kitchener, 1994), which define stages in terms of domain-specific structures like social perspective-taking or forms of relativism.

We have undertaken several studies of the reliability and validity of the LAS and its predecessors (Dawson-Tunik, 2004). We have examined inter-analyst agreement rates, compared scores obtained with the LAS with scores obtained with more conventional scoring systems, and examined scale characteristics with statistical modeling. Inter-analyst agreement rates have been high, 80% to 97% within half of a complexity level (Dawson-Tunik, 2004). Correspondences between the LAS and other developmental scoring systems are also high, consistently revealing agreement rates of 85% or greater within $\frac{1}{2}$ of a complexity level, although comparisons of the construct validity of the LAS and other systems have shown the LAS to be a more valid measure of cognitive performance (Dawson-Tunik, 2004). Employing Rasch scaling, which provides reliability estimates that are equivalent to Cronbach's alpha, we

have consistently calculated reliabilities over .95 (Dawson-Tunik et al., 2005). Overall, our research shows the LAS to be a valid and reliable general measure of intellectual development.

However, as explained above, developmental level as determined by the LAS is only one of the properties of leadership decision-making measured by the LDMA. The Lactical score is supplemented with 7 additional scales that provide insight into different aspects of the performance: perspective taking, perspective seeking, perspective coordination, collaborative capacity, contextual thinking, cognitive coherence, and decision-making process. Thus, the LDMA report provides detailed diagnostic insights into the learning needs of test-takers along 8 dimensions. Report feedback also includes developmentally targeted learning suggestions, reflective activities, and recommended learning resources such as books, websites, and videos. The LDMA can serve both formative and summative purposes, providing an objective measure of developmental complexity (the Lactical score), as well as a richly educative diagnostic report containing targeted learning materials. We now discuss the use of the LDMA as an embedded, formative assessment.

The Clear Impact Leadership Education Initiative: embedded assessments in action

Clear Impact Consulting Group, Inc. and Lectica, Inc. have collaborated in the development and delivery of a large-scale Leadership Effectiveness Initiative (LEI) for a major North American city government. This work stands as the best example to date of how LectaTests can be embedded in educational initiatives.

The Clear Impact LEI was conducted as part of an ambitious culture change being undertaken by a major North American city government, which included re-tooling their professional development offerings for managers. The LEI itself involved over 800 city leaders across the full range of management levels in the government. All participants were involved in 9

leadership development workshops, for a total of 40 hours of instructional time over 9 months. The Clear Impact curriculum focuses on foundational leadership skills such as reflection, perspective taking and seeking, emotional intelligence, contextual intelligence, and a variety of specific topics, such as high performance teams and change management. Integral principles inform the design of the curriculum, which includes the use of levels, lines, and states, as well as an orientation to the 4-quadrants. Classroom activities and take-home application-based learning assignments target both *what* leaders think, and *how* they think. This dual focus on the content and structure, as well as the explicit use of developmental levels, allowed for the integration of LectaTests into the curriculum as embedded assessments.

All participants in the LEI were asked to take up to 8 LectaTests as a part of the program. Each administration of a LectaTest was framed as a learning activity and thus embedded in the flow and delivery of the curriculum. At the same time each LectaTest provided the instructors with insights into the learning needs of their students, facilitating a process of data-driven dynamic curricular steering. The Lectual Decision Making Assessment (LDMA) was assigned 4 times, with pre-posts being required. The LDMA was used to structure classroom activities and whole group debriefs, where the assessment results were used to structure collective reflection and engagement. The Lectual Self-Understanding Assessment (LSUA) was assigned once and the results were used as a part of dyadic in-class discussion. The Lectual Leadership Reasoning Assessment (LLRA) was assigned once and used as a journaling exercise, which was then debriefed in class as part of a collective reflective practice. The Lectual Reflective Judgment Assessment (LRJA) was assigned once and used as a part of a small group assignment in which participants worked together to consider the assessment results. And finally, the Lectual Ethical Reasoning Assessment (LERA) was assigned once, and used as the focus of classroom discussion.

As this brief overview indicates, the relation between assessment, curriculum, and instruction in the Clear Impact LEI has been orchestrated so that LectaTests serve as learning catalysts for each student and scaffolds for classroom activities while providing the instructors with objective insights into their students' decision-making skills. LectaTests also provide data for research and evaluation purposes. In the next section, we describe how LDMA results were used to address two important research questions. First, did LEI participants who took more LectaTests, achieve greater Letical growth than those who took fewer? Second, did Letical growth predict improvements in decision-making behavior as reported by peers and direct reports?

Methods and results

In this section, we ask if embedding LectaTests in leadership programs increases the rate of individual growth. We define *embedding* as either using LectaTests to inform instruction or incorporating them into a curriculum (or coaching) to stimulate self-evaluation and reflection.

We present two types of evidence. First, we compare Letical growth across 8 program evaluations in which LectaTests were and were not embedded. Second, we examine how embedding LectaTests in Clear Impact's leadership program affected the growth of managers and their direct reports.

Program comparison

Here, we compare the average Letical growth that occurred in 5 programs in which LectaTests were *not* embedded with average Letical growth in 3 programs in which LectaTests were embedded.

Methods

Programs included in this comparison were delivered in a number of sectors, including business, government, and higher education. As shown in Table 1, individuals participating in these programs ranged in age from 18 to 65.

Table 1: *Characteristics of Programs*

<i>Program</i>	<i>Sector</i>	<i>N</i>	<i>Age Range</i>
Critical thinking-IC 2004, LDMA	Government	40	24-59
DM-1-AU 2010, LDMA	Law, medicine	28	31-64
Critical thinking 2010-MH, LRJA	Higher Ed	43	18-22
DM-3-NA 2012, LDMA	Higher Ed	24	19-32
DM-4-ST 2012, LDMA	Business	14	29-65
DM-2-IC 2005, LDMA	Government	32	24-55
Columbia University-ZV, 2012	Higher Ed	18	23-55
DM-Clear Impact 2013, LDMA	Government	512	25-65

In the majority of these programs, the Letical Decision Making Assessment (LDMA) was employed to evaluate program effectiveness. Participants in one program took the LRJA, an assessment of reflective judgment. The Letical Assessment System (LAS) was used to score responses for their developmental level, yielding scores from 10.5 to 12.3, which is typical in adult samples.

Participants in all programs completed at least two LectaTests, one before the start of the program and one after its completion. The length of programs varied from 1 to 12 months, with 40 to 60 hours of instruction. We calculated the average change in Lectical scores for participants in each program.

Results

As shown in Table 2, the results suggest that embedding LectaTests has a positive impact on development. Participants in programs in which LectaTests were embedded grew .21 of a level on average, whereas participants in programs without embedding grew .09 of a level on average¹. It is also interesting to note that hours of instruction and program duration did not appear to explain growth. For example, participants in a 60-hour, 6-month educational program without embedding grew only .06 of a level on average, while participants in a program of the same duration in which assessments were embedded grew .27 of a level on average. It is also useful to note that the gains in the Clear Impact sample occurred despite consistent perception on the part of leaders that organizational support for their participation was lacking, including having time to do action learning between sessions, and often having to miss sessions because of other organizational priorities. In addition, the participation of many of the leaders was not voluntary.

¹ Average annual growth for college students is .13 of a level.

Table 2: *Results From 9 Studies*

<i>Study</i>	<i>Embedded</i>	<i>Interval</i>	<i>Hours</i>	<i>Change</i>
Critical Thinking-IC 2004, LDMA	No	6 months	60	0.06
DM-1-AU 2010, LDMA	No	12 months	43	0.09
Critical Thinking 2010-MH, LRJA	No	12 months	42	0.13
DM-3-NA 2012, LDMA	No	1-5 months	40	0.03
DM-4-ST 2012, LDMA	No	6 months	40	0.15
DM-2-IC 2005, LDMA	Yes	6 months	60	0.27
Columbia University-ZV, 2012	Yes	4 months	40	0.18
Clear Impact 2013, LDMA	Yes	9 months	40	0.18

Clear Impact Consulting Group Study

Here we examine how the use of embedded assessments affected the growth and behavior of managers participating in the Clear Impact LEI program. This program is described in detail, above.

Methods

Of the managers who had finished the LEI program at the time of this analysis, 161 had completed both their pre and post LDMA. All managers who completed their post program

LDMA also filled out a “360” type survey. In this survey, participants were asked to comment on behavioral change in themselves and their peers, supervisors, and direct reports during the LEI program. This survey contained several questions about aspects of decision-making behavior, such as the following:

- Since the beginning of this program, I more frequently take the perspectives of others into account.
- Since the beginning of this program, my peers more frequently seek the perspectives of others.
- Since the beginning of this program, my supervisor more frequently asks questions that invite people to think and reflect.

Of the 161 managers who completed pre and post LDMA, 10 of them had two or more direct reports who had also completed pre/post LDMA and “360” surveys at the time of this analysis, making it possible to examine the relation between the 360 results of subordinates and peers and their Lectical growth.

Results

Our analysis produced four key findings, summarized in Table 3 and Figure 1.

Table 3: *Four Key Findings*

<i>Finding</i>	<i>N</i>	<i>R</i>	<i>P</i>
Supervisors who took more LectaTests experienced more Letical growth	161	0.19	0.01
Supervisors with greater Letical growth received higher 360 scores from direct reports	9	0.59	0.09
The direct reports of supervisors who received higher 360 scores experienced more Letical growth	10	0.60	0.07
Direct reports who experienced more Letical growth received higher 360 scores from peers.	10	0.46	0.27

First, of the 161 participants who completed both their pre and post LDMA, those who completed more of the 8 assigned LectaTests demonstrated greater Letical growth over the course of the program.

The next three findings concern the relationship between supervisors' personal growth, the growth of their direct reports, their direct reports' evaluation of them, and their direct reports' evaluation of their peers. The small sample size made it virtually impossible to generate

statistically significant ($p < .05$) results. Although the correlations (Rs) are large, suggesting moderate to large effects, none of the analyses generated statistically significant results.

Consequently, we view these results as suggestive and tentative.

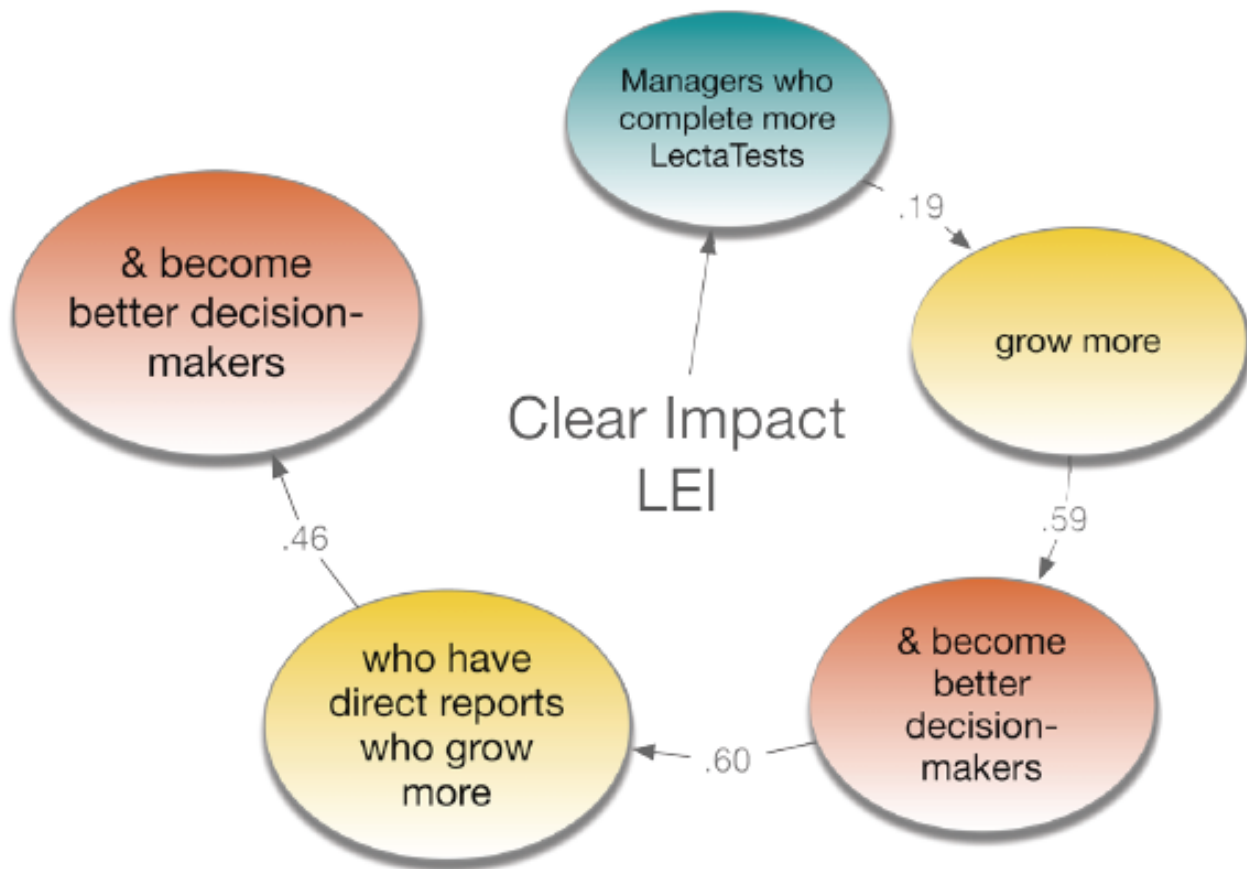
The second finding is that supervisors who experienced greater Lectical growth received higher “360” scores from their direct reports, suggesting that developing managers’ reasoning about decision-making can change their behavior. The third finding is that the direct reports of supervisors who received higher 360 scores experienced greater Lectical growth themselves.

This apparent connection between leaders’ behavior and the growth of their subordinates suggests that a leaders’ ability to model and demonstrate more effective decision making can support the development of this skill in their subordinates. The final finding is that direct reports who experienced more Lectical growth received higher 360 scores from their peers.

This result reinforces the connection between Lectical growth and perceived behavior change.

Overall, results suggest that managers who participated more fully in the LEI not only learned new ideas and developed greater capacity for complex thought, but also learned actionable decision making skills that are being applied in the workplace.

Figure 1: *LEI growth spiral*



Discussion: virtuous cycles of learning

We began by considering a broad philosophy of learning that suggests educational initiatives should be structured to allow for iterative cycles of action, feedback, and reflection. Creating educational environments that promote these *virtuous cycles of learning* requires building

assessment infrastructures that can be used for both summative and formative purposes—objective measures that also serve as educational diagnostics and learning aids. We then described the research and theory behind LectaTests, which are designed to play both summative and formative roles. The Clear Impact LEI was presented as an example of how LectaTests can be embedded as a part of classroom practices in leadership education. When embedded in this way LectaTests promote development on the part of the learner, aid the educator in the delivery of curricula, and serve as objective program evaluation tools. Finally, we presented findings that demonstrate the potential power of embedded assessments for promoting cognitive growth and in changing the behaviors of leaders on the ground. These findings support the broad philosophy of learning introduced at the outset, and suggest that learning can be catalyzed through the use of embedded, formative assessments.

While it seems clear that embedding LectaTests makes a difference in educational contexts, there is a great deal we don't know about how and why this is. We are just beginning to explore the possibilities for action-research using LectaTests, and can already see important directions for future research. What are the best practices for embedding LectaTests in various contexts, and which practices promote the most robust growth? Do some approaches promote continued growth over the long run, with diagnostic assessments contributing to continued learning beyond the classroom? What about transfer—does the growth resulting from the use of embedded assessments support growth in areas not directly targeted by the original educational intervention? And finally, what are the most cost effective ways to use of LectaTests in different organizational contexts? These and other questions will drive future research about the use of LectaTests as formative assessments. Answers to these questions will help shape the future of innovation in educational assessment technologies.

Sources:

- Baldwin, J. M. (1906). *Mental development in the child and the race: Methods and processes*. New York: The Macmillan Company.
- Bhaskar, Roy. (1991). *Dialectic: the pulse of freedom*. New York: Verso.
- Brainerd, C. J. (1993). Cognitive development is abrupt (but not stage-like). In H. Thomas & A. Lohaus (Eds.), *Modeling growth and individual differences in spatial tasks* (Vol. 58(9), pp. 170-190).
- Colby, Anne, & Kohlberg, Lawrence. (1987). *The measurement of moral judgment: Standard issue scoring manual* (Vol. 1). New York: Cambridge University Press.
- Commons, Michael L., Straughn, J., Meaney, Maryellen, Johnstone, Jonalu, Weaver, Julia H., Lichtenbaum, Erica, . . . Rodriquez, Joseph. (1995, November). *Measuring stage of development across domains: A universal scoring scheme*. Paper presented at the Annual meeting of the Association for Moral Education, New York.
- Cook-Greuter, Susanne Regina. (1999). *Postautonomous ego development: A study of its nature and measurement*. (Ed.D. 9933122), Harvard University, United States -- Massachusetts. Retrieved from <http://search.proquest.com.ezp-prod1.hul.harvard.edu/docview/304503009?accountid=11311> Dissertations & Theses @ Harvard University; ProQuest Dissertations & Theses (PQDT) database.
- Dawson, Theo L. (2002). A comparison of three developmental stage scoring systems. *Journal of Applied Measurement*, 3, 146-189.
- Dawson, Theo L. (2010). The Lectical™ Assessment System. Retrieved March 30, 2010, from <http://lectica.info>
- Dawson, Theo L., & Stein, Zachary. (2004). National Leadership Study results. Hatfield, MA: Developmental Testing Service, LLC.
- Dawson, Theo L., & Stein, Zachary. (2006). Decision-making curriculum. Results of the pre- and post-instruction developmental assessments. Northampton, MA: Developmental Testing Service.
- Dawson, Theo L., & Stein, Zachary. (2011). We are all learning here: Cycles of research and application in adult development. In C. Hoare (Ed.), *The Oxford Handbook of reciprocal adult development and learning* (pp. 447-460). New York: Oxford.
- Dawson, Theo L., & Wilson, Mark. (2004). The LAAS: A computerized developmental scoring system for small- and large-scale assessments. *Educational Assessment*, 9, 153-191.

- Dawson-Tunik, Theo L. (2004). "A good education is..." The development of evaluative thought across the life-span. *Genetic, Social, and General Psychology Monographs*, 130(1), 4-112.
- Dawson-Tunik, Theo L. (2006). The meaning and measurement of conceptual development in adulthood. In C. Hoare (Ed.), *The intersection of adult development and learning: A handbook of theory, research, and practice* (pp. 433-454). London: Oxford.
- Dawson-Tunik, Theo L., Commons, Michael L., Wilson, Mark, & Fischer, Kurt W. (2005). The shape of development. *The European Journal of Developmental Psychology*, 2(2), 163-196.
- Fischer, Kurt W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87, 477-531.
- Fischer, Kurt W. , & Dawson-Tunik, Theo L. (2006, April). *Applying cognitive science to curricular design and analysis of concept learning & development*. Paper presented at the Learning and the Brain, Cambridge, MA.
- Fischer, Kurt W., & Bidell, T. R. (2006). Dynamic development of action, thought, and emotion. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Theoretical models of human development* (6 ed., Vol. 1, pp. 313-399). New York: Wiley.
- Forrester, Jay. (1964). *Industrial dynamics*. Cambridge, MA: MIT Press.
- Habermas, Jurgen. (1984). *The theory of communicative action: reason and the rationalization of society* (T. McCarthy, Trans. Vol. 1). Boston: Beacon Press.
- Kegan, Robert. (1982). *The evolving self: Problem and process in human development*. Cambridge, Massachusetts: Harvard University Press.
- Kegan, Robert, & Lahey, L. (2002). *How the way we talk can change the way we work: Seven languages for transformation*. San Francisco: Jossey-Bass.
- King, P. M., & Kitchener, K. S. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults*. San Francisco: Jossey Bass.
- Kohlberg, Lawrence. (1969). Stage and sequence: The cognitive-developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347-480). Chicago: Rand McNally.
- Loevinger, Jane. (1966). The meaning and measurement of ego development. *American Psychologist*, 21(3), 195-206.
- Peirce, C.S. (1865). On the logic of science, Harvard Lectures of 1865. In P. e. project (Ed.), *Writings of Charles S. Peirce: a chronological edition* (Vol. 1, pp. 162-268). Bloomington, IN: Indiana University Press.
- Perry, W. G. (1970). *Forms of intellectual and ethical development in the college years*. New York: Holt, Rinehart, & Winston.

Piaget, Jean. (1952). *The origins of intelligence in children* (M. Cook, Trans.). New York: International Universities Press, Inc.

Piaget, Jean. (1971). *Biology and knowledge: An essay on the relations between organic regulations and cognitive processes*. Chicago, IL: U. Chicago Press.

Waddington, C. H. (1966). *New patterns in genetics and development*. New York: Columbia University Press.

Wilber, K. (1999). *Integral Psychology*. Boston: Shambhala.