WILEY

Go for the gold: Hurdles and winning strategies for conducting longitudinal and nonlinear research in undergraduate leader development

Jessica B. Diaz[1] | **Rebecca J. Reichard**[1] | **Ronald E. Riggio**[2]

[1]Division of Behavioral and Social Sciences, Claremont Graduate University, 150 E. 10th Street, Claremont, CA 91711

[2]Kravis Leadership Institute, Claremont McKenna College, 888 Columbia Avenue, Claremont, CA 91711

Correspondence
Jessica B. Diaz, Division of Behavioral and Social Sciences, Claremont Graduate University, 150 E. 10th Street, Claremont, CA 91711.
Email: jessica.diaz@cgu.edu

Abstract
This article makes a case for longitudinal and nonlinear methods when researching or evaluating student leadership development. After a primer on longitudinal methodology, barriers and aligned solutions to methodological challenges are presented.

Advancing research and program evaluation in student leadership development requires methods and analytical approaches that parallel the natural development process. Leader development (LD), like all forms of human development, occurs over a long period and is a process, albeit not a linear one. The same is true for leadership development, which focuses on developing a group's collective capacity for leadership. Both consists of periods of advancement, decline, and times where development may stagnate. Here is an example.

> After having much success in his small rural high school and feeling confident in his leadership, Jamal quickly enrolls in the undergraduate leadership program during his first year (Time 1). While completing the Introduction to Leadership course, he is impressed with the leadership knowledge and skills that the senior student course mentors have acquired. By the end of the semester, Jamal realizes that he has much to learn about leadership and feels slightly less confident (Time 2). Eager to learn more, he signs up for the leadership sequence's second course and seeks out an elected position as Secretary/Treasurer in a campus club. By the end of his second year, Jamal regains some of his confidence as a leader (Time 3). In his third year, he serves as President in the campus club and completes a course related to culture and leadership. The club goes through some scrutiny regarding how it spent funds

in previous years, and as the current president, campus administration goes to Jamal for answers. He struggles emotionally and seeks out coaching and mentoring from leadership faculty. At the same time, in his culture and leadership class, Jamal is exposed to perspectives and ideas that counter his own deeply held assumptions. Thus, as a Junior, Jamal deeply questions who he is as a leader (Time 4). In his senior year, Jamal completes the capstone leadership course, serves as a Senior course mentor in the introductory course, and continues his service as Club President. Throughout these experiences, Jamal reflects on the lessons from his junior year. He thrives as a mentor and leader. Jamal applies what he has learned and emerges from his undergraduate experience an even more confident leader (Time 5).

Understanding the trajectory of Jamal's leadership journey requires that researchers employ longitudinal methods, a winning methodology for studying and evaluating LD. However, just as Olympic athletes must overcome obstacles to reach the podium, going for gold in LD research presents several hurdles for leadership educators. Some of the significant limitations are understanding what longitudinal methods include, knowing what to measure and how to measure it, dealing with attrition and missing data, and managing and analyzing longitudinal data. Given these hurdles, it is not surprising that most LD studies are not *true* longitudinal studies (Day, 2011). Most are cross-sectional (i.e., examining data at one point), and extant longitudinal studies tend to be short-term in their focus, looking at changes over only a week, month, or semester (Riggio & Mumford, 2011). Cross-sectional or short-term studies require less time, fewer resources, and more straightforward analyses. Yet, they fall short of the gold standard for a true longitudinal study. Going for the gold requires three or more waves of data collection, tracking developmental outcomes over time with valid measures, and using realistic time periods (i.e., months or years, rather than days) to allow development to occur (Day, 2011). This article aims to empower leadership educators with the knowledge to design and implement longitudinal methods. After describing what longitudinal methods are, we detail specific hurdles and how to jump over them toward victory!

NEED FOR LONGITUDINAL AND NON-LINEAR METHODS

As we see with Jamal, LD is a long-term process. Experienced leadership educators understand that there is no "magic pill" that helps students develop their capacity to lead effectively. Instead, LD occurs over the long term and involves a deep change in identity, schemas, internal structures, and conceptual frames (Day & Dragoni, 2015). It arises from an ongoing series of challenging curricular and co-curricular experiences that require hard work, deep engagement, and self-reflection (Murphy & Johnson, 2011; Walker & Reichard, 2020). If we take a snapshot of students' leadership outcomes at one point—say at the end of their undergraduate career (e.g., Time 5 for Jamal)—we fail to capture the ebb and flow of the longitudinal development process. As a result, we cannot draw any firm understanding of the development process or its key drivers. Moreover, we cannot even conclude that any LD occurred if we only observe snapshots at Time 1 and Time 5. We know from Jamal's story that his developmental trajectory is far more complex and non-linear.

In addition to the hurdle of non-linearity, not every individual grows at the same rate or in the same way. Most easy analytical processes for longitudinal data produce results that provide researchers with a single relationship between variables for all undergraduate students in a sample. Looking at differences between students, a researcher can draw general conclusions about growth in LD over time. Although this might provide

important insight into the relationships in question, particularly the non-linearity inherent in growth trajectories, it ignores how between-student differences may alter these trajectories. For example, in this analysis, Jamal, who experienced considerable growth in his LD over his undergraduate career, might be offset by other students who did not experience much change. As a result, the researcher might conclude that the LD program does not lead to development, although Jamal (and potentially others) benefitted from participating.

Exemplar LD researchers must deliberately consider how to account for both non-linearity and these individual differences in growth. Such considerations require more sophisticated analyses that extend beyond the basic approaches taught in most graduate methods courses.

For example, in their study of about 1000 undergraduates engaged in team action learning programs designed to improve participants' leadership, Day and Sin (2011) examined how individual leader identity impacts the growth trajectory of participants. A linear examination of LD over time was, on average, negative, suggesting a slight *decrease* throughout the semester. After modeling the average development non-linearly, the authors found a quadratic relationship (initial decrease followed by eventual rebound) better explained the average development trend. Disaggregating their results into individual trajectories, they found that students with stronger initial leader identities demonstrated greater development in leader effectiveness over time. Additionally, they found significantly different trajectory shapes based on participants' level of learning goal orientation (i.e., approaching challenges as an opportunity to grow). Leadership ratings for individuals with lower learning goal orientation decreased linearly across the four time periods. In comparison, leadership ratings for participants with higher learning goal orientation stabilized and then rebounded by the end of the semester. Examining average trajectories across all study participants would have missed these between-person effects.

WHAT ARE LONGITUDINAL METHODS?

Although the term longitudinal is generally applied to any study where data are collected on variables of interest over time, a review of existing longitudinal research reveals significant disparities in the design, conduct, and analysis procedures of longitudinal research. At the core, longitudinal analysis involves the study of individuals as they exist and change over time (Nesselroade & Baltes, 1979) and involves some element of repeated measurement. However, what researchers mean by longitudinal studies varies and is often misrepresented (Ployhart & Vandenberg, 2010). To help researchers identify and conduct longitudinal research, there are three principal methodological elements: (a) three or more data collection time points, (b) a focal outcome that changes systematically over time, and (c) an appropriate metric for modeling the element of time (Day, 2011; Singer & Willett, 2003).

Returning to the Olympic analogy, we offer a three-level framework of LD research design rooted in these three methodological elements of longitudinal research. From bronze to gold, each approach increases in utility and complexity.

Cross-sectional design (bronze)

Longitudinal methods track the same LD outcome over time. Thus, our last-place finisher—cross-sectional design—does not qualify as longitudinal. Cross-sectional studies refer to variables captured at a single point in time and do not include any of the three elements of a longitudinal design. Applied to the context of LD, this might look like comparing

the LD outcome of leader identity of first-year undergraduates to that of seniors as a proxy for determining whether leader identity develops throughout an individual's undergraduate experience. Although findings from such a study can infer associations between the year in school and leader identity, it is impossible to extrapolate results related to change, as any observed difference can be attributed to a host of alternative factors (Singer & Willett, 2003). Put simply, cross-sectional studies are analogous to a coach comparing a novice athlete's response to setbacks with that of an Olympic veteran to conclude how athletic resilience develops over time. Both aim to study inherently dynamic change processes in static form (Ployhart & Vandenberg, 2010).

Quasi-longitudinal (silver)

Coming in second place are quasi-longitudinal research designs (Day, 2011). Analogous to the term quasi-experimental (Cook et al., 1979), quasi-longitudinal studies generally include two of the three elements of a longitudinal design. For example, Reichard et al. (2021) examined the change in the LD outcome of implicit leadership theory constellations before and after the student's first year of college. They found that students' definitions of the traits and characteristics of an effective leader shifted from a vague and varied grouping of traits toward a more cohesive and organized understanding of leadership. Interestingly, the researchers identified that students completing leadership coursework were more likely to experience this shift in their thinking about leadership. Although this study focused on an LD outcome that can change over time (i.e., implicit leadership theories) and used a reasonable measure of time (i.e., the beginning and end of the first year of college), it fails to capture what happens in between the measurement points.

Quasi-longitudinal research can also span many years, as we observe in the ongoing Fullerton Longitudinal Study (Oliver et al., 2011; Reichard et al., 2011). In 1979, researchers set out to examine the long-term development of 130 one-year-olds, with particular scholars examining leadership outcomes in adulthood. Current published findings reflect the quasi-longitudinal design examining the developmental roots of leadership. For example, Guerin et al. (2011) found that adult social skills fully mediate (i.e., explain) the relationship between adolescent extraversion and adult leadership. As researchers assess additional waves of the same leadership outcomes, this study will allow true longitudinal analysis.

Quasi-longitudinal designs limit the conclusions researchers can draw. For example, two timepoints can only be modeled linearly (e.g., preference for a particular implicit leadership theory constellation can only go up or down between two-time points). In reality, college students' intra-individual implicit leadership theories may vary significantly throughout their first year. Additionally, quasi-longitudinal designs confound change and error measurement. It is impossible to eliminate the possibility that the measurement used had differential effects at each timepoint (Singer & Willett, 2003). Although two-time points are better than one, it does not tell us much about the LD process.

True longitudinal (gold)

The gold medal winner is "true longitudinal studies" (Day, 2011, p. 562)—three or more waves of data, track an LD outcome over time, and adopt a sensible time metric. Take the proximal LD outcome of leader identity, for example. Leader identity reflects the extent to which being a leader is central to one's self-concept (Day & Dragoni, 2015). Multiple *true* longitudinal studies have shown that trajectories of leader identity ebb and flow over time. For example, Middleton et al. (2019) assessed the leader identity of 39 executives

12 times throughout a 5-month executive LD program. Data indicated a significant positive linear increase in leader identity over time, with the growth slowing toward the latter half of the program. Miscenko et al. (2017), studying leader identity in 98 business students throughout a 7-week leadership course, found that leader identity develops in a *J*-shaped pattern, a decrease followed by an increase. In both studies, researchers examined the relationship between other individual variables and leader identity changes (i.e., learning goal orientation and leadership skills, respectively), shedding light on factors that influence the developmental trajectory. True longitudinal studies win gold because they examine LD over time, accounting for its non-linear and dynamic nature.

Another gold-level example is research being conducted on United States Military Academy cadets at West Point since 2001. Specifically, Harms et al. (2011) examined how various personality traits impacted a cadet's rate of leader development over a multi-year period. They found some traits (e.g., skeptical) negatively impact leader development over time while others (e.g., bold) have a catalyzing effect. A study of this magnitude presents several challenges, many of which we discuss below. Nevertheless, it is an example of the gold standard for what longitudinal LD looks like.

Although Harms et al. capture these differences quantitatively, we would be remiss without briefly mentioning the benefit of a mixed-method approach. Adding a qualitative data source, such as interviews, to a gold-level quantitative study can deepen understanding of how trajectories may differ for individuals (see Article 10 for more on mixed methods).

GAP BETWEEN HOW LEADERS DEVELOP AND USE OF LONGITUDINAL METHODS: WHY AND WHAT CAN BE DONE?

As noted, studying LD must consider the role of time. Unfortunately, time often gets overlooked in the design of research on LD in one of three ways—timespan, frequency, and analysis of measures. First, it is challenging to determine how long it takes leaders to develop. Different leadership skills or mindsets change at different rates, requiring different time lengths for investigation (Day et al., 2021). Although, lengthy studies pose their challenges, truncated studies may threaten the validity of the findings, particularly for non-significant results. In this case, it is nearly impossible to distinguish the cause of such non-significant results. Was there a genuine lack of growth or an insufficient amount of time for change to occur?

Second, there is often a mismatch between the natural development process and the design and analysis methods used to study it. Study designs employing too few measurement points may fail to capture curvilinear (i.e., *J*-shaped) growth trajectories (Day et al., 2021). For example, if researchers only captured Jamal's leadership in his junior year, they would deduce a negative development trajectory when his LD was far more complex.

Researchers must also be mindful of how LD occurs over time from a statistical standpoint. Simple analyses that model the average change over time (e.g., linear regression) are not always best suited for LD research. They limit our understanding of growth to either increases or decreases, never allowing for the possibility of both. This method also treats deviations in individual trajectories as model error.

HURDLES AND WINNING STRATEGIES TO LONGITUDINAL EXAMINATION OF STUDENT LEADER DEVELOPMENT

Considering both the importance and difficulty of longitudinal and non-linear approaches to studying LD, the remaining sections discuss hurdles and winning strategies to

successfully designing and analyzing gold medal longitudinal research and program evaluation. We specifically highlight four categories of challenges: what to measure, how to measure it, how to monitor it, and how to analyze it. For each hurdle, we offer practical solutions from both the literature and our collective experience with longitudinal research.

Hurdle 1: What LD outcome to measure—the criterion problem

The first hurdle in any student LD study is determining which LD outcomes are of primary interest. What does a successful LD "journey" look like? Does it mean that students have mastered a particular leadership skill, are more confident leaders, or emerge in leadership roles? Although tempting, measuring too many variables can hinder our ability to detect whether development occurred. Instead, longitudinal research requires precision and parsimony about which leadership variables are expected to change as well as how and why they might do so (Ployhart & Vandenberg, 2010). Theory, prior research, and the university or program of interest should guide this determination.

Researchers and evaluators must consider two additional areas when determining what to measure. The first is the level of analysis, which refers to whose development we measure over time. There are two commonly used levels of analysis in LD research. As we discussed earlier in this article, the first involves determining the average change or change trajectory for an entire group of students. Alternatively, researchers may be interested in intra-individual change—focusing on the individual changes over time.

Finally, choosing LD outcomes involves deciding whether to focus on outcomes that develop more quickly or those that develop later in an individual's trajectory. In their seminal review, Day and Dragoni (2015) integrated time into a longitudinal framework of LD outcomes, characterizing LD outcomes according to the speed at which they develop. Proximal indicators (relevant in shorter time frames) include leader self-views (i.e., leader efficacy, leader identity, and self-awareness), specific leadership knowledge and skills, and positive leadership behaviors. Other proximal LD outcomes may include leader meaningfulness and implicit leadership theories. More distal outcomes (relevant for long-term designs) include advanced and complex leadership skills, rates of career advancement, promotions, etc.

Clarity on what to measure begins with a strong research question. Ployhart and Vandenberg (2010) distinguish between descriptive versus explanatory research, a differentiation that proves helpful when designing research questions for longitudinal studies. Descriptive research describes how focal variables change over time, while explanatory research aims to uncover the causal processes driving that change.

Hurdle 2: How to measure the chosen LD outcome—rigor

A second hurdle to effective longitudinal (LD) research relates to its methodological rigor. Although few researchers set out to design non-rigorous research, practical considerations often lead to decisions or sacrifices that have important implications for the study design and findings. Relevant to LD research are four threats to rigor: (a) mistaking evaluation for research, (b) lack of an equivalent comparison group, (c) inadequate measures, and (d) inappropriate time spacing.

Evaluation v. research

Evaluations of LD programs tend to focus on efficacy regarding how participants rate the program and any growth in leadership that participants attributed to their participation.

However, it can be easy to mistake evaluation for research, a pitfall discussed extensively in article two. A well-known program evaluation model (Kirkpatrick, 1994) helps make this distinction. *Reaction criteria*, which focus on participants' enjoyment and attitudes toward the program, are commonly used to evaluate LD programs, but they are extremely limited. It tells us what participants think about the LD program, but not whether any change occurred. The next level, *knowledge criteria*, assess whether participants learned anything about leadership and might be measured through an exam. While this makes sense for college-level assessment, it still does not relate to behavioral change. The third level of evaluation is where the evaluation program moves more toward research. *Behavioral criteria* assess actual behavioral change in participants' leader behaviors, such as attaining leadership positions or displaying learned leadership skills. In short, LD research focuses on generating new understandings about how and why individuals develop over time, while evaluation focuses on assessing program effectiveness. Despite their differences, both are critical players in leadership education.

Equivalent comparison group

Undergraduate LD programs intend to provide some growth in leadership potential apart from naturally occurring maturation processes. To adequately assess growth, it is important to have some comparison group of non-participants. However, because most students opt into LD programs, either voluntarily or through participation in an academic leadership program, there may be motivational differences between those who choose LD program participation and those who do not. These differences can make it challenging to compare developmental differences between the two groups. Attention must be given to ensuring that comparison groups are equivalent, meaning they are similar in terms of demographics, attitudes, etc., to those whose development is being studied. Strategies include randomly assigning only half of the participants to the LD program while offering the non-participating comparison group the LD program later (Article one discusses the creative use of program waitlists). Another strategy would be to measure potential confounding variables that might influence LD outcomes and attempt to control for them statistically.

Adequate measures of LD outcomes

As in any study, it is necessary that the measures used be both valid (i.e., they measure the LD outcome you want to measure) and reliable (i.e., they measure the same way across participants and time). Rather than making up questions, LD outcomes should be assessed with previously validated, multi-item scales where information on validity and reliability (e.g., factor analysis results, Cronbach's Alpha, etc.) is available and robust. It is also important that there are enough items with enough answer choices to capture a wide range (i.e., variance) of student responses. Even for high-quality measures, longitudinal research presents additional measurement challenges. Most notably, it is essential to check whether participants interpret your measures the same way at different time points. This process is often called measurement invariance and can be analyzed by trained statisticians.

Inappropriate time spacing

The number and spacing of repeated observations pose a final methodological challenge. Each additional wave of data collection increases the power and reliability of one's

findings, yet practical limitations often constrain the number of possible data collection time points. The determinants for the number of observations ought to depend on the research question and the hypothesized change trajectory of the specific outcomes under investigation. That is, it is essential to have enough time points to capture the specific form of change in the LD outcome you are investigating (Ployhart & Vandenberg, 2010). Regarding spacing, determining the length of time between collection intervals once again depends on both the variables and the population of interest. For example, in one LD evaluation of undergraduate business majors, work supervisors rated the now-alumni's "leadership potential" 2–3 years post-graduation (Riggio et al., 2003).

Hurdle 3: How to monitor the LD outcome—attrition

Multiple waves of data collection are paramount to ensuring methodological rigor in any undergraduate LD study. However, from a practical perspective, this presents a host of challenges, chief among them attrition, post-graduation tracking, and control monitoring. Attrition refers to the loss of study participants at some point during the data collection process, the risk being loss of statistical power resulting from a smaller sample size. However, when attrition is selective, and the causes are systematic, the impact can threaten the validity of findings (Crano et al., 2014). That is, students at later time points must be representative of those who eventually ceased participation.

Although evidence-based research on avoiding attrition in longitudinal studies is scarce, we offer several winning strategies. The first suggested strategy is to develop an a priori theory about what might cause attrition in your sample and then measure this during your first wave of data collection (Ployhart & Vandenberg, 2010). For example, suppose you believe that students may drop out because they are not sufficiently motivated to develop their leadership. In that case, you might use a measure of "motivation to develop" to see if this accounts for the attrition. It is also important to simply acknowledge that attrition is inevitable in longitudinal studies and plan to start with a much larger sample than ultimately desired. A final strategy for combatting attrition is to collect locator information at the onset of the study. This approach involves having each student nominate three locators (e.g., parent, sibling) who will always know where they are (Hill et al., 2016).

Hurdle 4: How to analyze—advanced statistical procedures required

Beyond methodology, longitudinal and non-linear analyses of LD present a host of issues, including determining which statistical approaches are best for different types of questions. As we have established, we rarely expect LD growth trajectories to be linear. Curvilinear relationships may be more expected. Alternatively, abrupt directional changes may occur in an otherwise linear growth trajectory (i.e., breaking points; Ployhart & Vandenberg, 2010) because of planned interventions such as LD programming or unplanned disruptions such as the recent COVID-19 pandemic.

Even when we expect growth to be linear, the nature of longitudinal data tends to violate the general linear model assumptions. For example, to ensure unbiased estimates, the general linear model requires that errors be uncorrelated (i.e., the degree to which your specified model fits your data at time one must be independent of how it fits at time two). In longitudinal research, we *would* expect error correlation given that variables are being measured repeatedly in the same individuals.

Finally, part of modeling development is determining one's approach to examining time as a predictor of change. Although beyond the scope of this article, multiple approaches (i.e., polynomials, orthogonal polynomials, scaled time metrics) exist for modeling time. There are now multiple ways to analyze longitudinal and non-linear development trajectories, including latent growth curve modeling, coupled latent growth curve modeling, cross-lagged modeling, and latent change score modeling. Many ongoing longitudinal studies enlist the aid of statistical experts who are well-versed in these advanced statistical methods—a winning strategy for overcoming the analysis hurdle.

CONCLUSION

Conducting longitudinal research on LD is not easy. The approaches that best capture the complexity of LD are often the most challenging from both a methods and analysis perspective. Many scholars feel pressure to "publish or perish," which leads some of them to use more straightforward and expedient research designs (e.g., cross-sectional research). Practitioners offering LD programs are under a similar time crunch to assess whether their methods and programs are indeed working, which can sometimes lead toward evaluation in place of research.

Longitudinal methods are fraught with many hurdles that seemingly rival the contribution of their findings while spending years in pursuit of an answer to a focused research question runs counter to our human need for speed. Yet if our goal is to capture the undeniably complex student development process, we must acknowledge that many traditional research approaches may not do justice to the processes we aim to explain. The best research uses the best available designs and methods, and for studying LD over time, true longitudinal research will help you go for the gold.

REFERENCES

Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings* (Vol. 351). Houghton Mifflin.

Crano, W. D., Brewer, M. B., & Lac, A. (2014). *Principles and methods of social research*. Routledge.

Day, D. V. (2011). Integrative perspectives on longitudinal investigations of leader development: From childhood through adulthood. *The Leadership Quarterly*, *22*(3), 561–571. https://doi.org/10.1016/j.leaqua.2011.04.012

Day, D. V., & Dragoni, L. (2015). Leadership development: An outcome-oriented review based on time and levels of analyses. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*, 133–156. https://doi.org/10.1146/annurev-orgpsych-032414-111328

Day, D. V., Riggio, R. E., Tan, S. J., & Conger, J. A. (2021). Advancing the science of 21st-century leadership development: Theory, research, and practice. *The Leadership Quarterly*, *32*(5), 101557. https://doi.org/10.1016/j.leaqua.2021.101557

Day, D. V., & Sin, H.-P. (2011). Longitudinal tests of an integrative model of leader development: Charting and understanding developmental trajectories. *The Leadership Quarterly*, *22*(3), 545–560. https://doi.org/10.1016/j.leaqua.2011.04.011

Guerin, D. W., Oliver, P. H., Gottfried, A. W., Gottfried, A. E., Reichard, R. J., & Riggio, R. E. (2011). Childhood and adolescent antecedents of social skills and leadership potential in adulthood: Temperamental approach/withdrawal and extraversion. *The Leadership Quarterly*, *22*(3), 482–494. https://doi.org/10.1016/j.leaqua.2011.04.006

Harms, P. D., Spain, S. M., & Hannah, S. T. (2011). Leader development and the dark side of personality. *The Leadership Quarterly*, *22*(3), 495–509. https://doi.org/10.1016/j.leaqua.2011.04.007

Hill, K. G., Woodward, D., Woelfel, T., Hawkins, J. D., & Green, S. (2016). Planning for long-term follow-up: Strategies learned from longitudinal studies. *Prevention Science*, *17*(7), 806–818. https://doi.org/10.1007/s11121-015-0610-7

Kirkpatrick, D. L. (1994). *Evaluating training programs: The four levels*. Berrett-Koehler.

Middleton, E. D., Walker, D. O., & Reichard, R. J. (2019). Developmental trajectories of leader identity: Role of learning goal orientation. *Journal of Leadership & Organizational Studies*, *26*(4), 495–509. https://doi.org/10.1177/1548051818781818

Miscenko, D., Guenter, H., & Day, D. V. (2017). Am I a leader? Examining leader identity development over time. *The Leadership Quarterly*, *28*, 605–620. https://doi.org/10.1016/j.leaqua.2017.01.004

Murphy, S. E., & Johnson, S. K. (2011). The benefits of a long-lens approach to leader development: Understanding the seeds of leadership. *The Leadership Quarterly*, *22*(3), 459–470. https://doi.org/10.1016/j.leaqua.2011.04.004

Nesselroade, J. R., & Baltes, P. B. (1979). *Longitudinal research in the study of behavior and development*. Academic Press.

Oliver, P. H., Gottfried, A. W., Guerin, D. W., Gottfried, A. E., Reichard, R. J., & Riggio, R. E. (2011). Adolescent family environmental antecedents to transformational leadership potential: A longitudinal mediational analysis. *The Leadership Quarterly*, *22*(3), 535–544. https://doi.org/10.1016/j.leaqua.2011.04.010

Ployhart, R. E., & Vandenberg, R. J. (2010). Longitudinal research: The theory, design, and analysis of change. *Journal of Management*, *36*(1), 94–120. https://doi.org/10.1177/0149206309352110

Reichard, R. J., Ellis, B., Powers, K. W., Walker, D. O., & Priest, K. (2021). Changes in implicit leadership theory constellations throughout a developmental experience. *Journal of Leadership Education*, *20*(4), 1–18. https://doi.org/10.12806/v20/i4/r12

Reichard, R. J., Riggio, R. E., Guerin, D. W., Oliver, P. H., Gottfried, A. W., & Gottfried, A. E. (2011). A longitudinal analysis of relationships between adolescent personality and intelligence with adult leader emergence and transformational leadership. *The Leadership Quarterly*, *22*(3), 471–481. https://doi.org/10.1016/j.leaqua.2011.04.005

Riggio, R. E., Mayes, B. T., & Schleicher, D. J. (2003). Using assessment center methods for measuring undergraduate business student outcomes. *Journal of Management Inquiry*, *12*(1), 68–78. https://doi.org/10.1177/1056492602250519

Riggio, R. E., & Mumford, M. D. (2011). Introduction to the special issue: Longitudinal studies of leadership development. *The Leadership Quarterly*, *22*(3), 453–456. https://doi.org/10.1016/j.leaqua.2011.04.002

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press.

Walker, D. O., & Reichard, R. J. (2020). On purpose: Leader self-development and the meaning of purposeful engagement. *Journal of Leadership Studies*, *14*(1), 26–38. https://doi.org/10.1002/jls.21680

AUTHOR BIOGRAPHIES

Jessica B. Diaz is an Assistant Professor in the Division of Behavioral and Organizational Science at Claremont Graduate University. There, she serves as Director of the Human Resource Management program and teaches the required year-long statistics sequence to all incoming MA and PhD students in psychology and evaluation. Her research and consulting work focus on leadership and the employee experience, including psychological safety, leader development, employee engagement, and diversity, equity, and inclusion.

Rebecca J. Reichard is an Associate Professor of Organizational Behavior at Claremont Graduate University. Her research focuses on leadership evaluation, assessment, and development. She is a founder and the director of LeAD Labs, which researches and applies evidence-based approaches to leader development.

Ronald E. Riggio is the Henry R. Kravis Professor of Leadership and Organizational Psychology at Claremont McKenna College. Dr. Riggio is a social/personality psychologist and leadership scholar with two dozen authored or edited books and more than 200 articles/book chapters. His research interests are in leadership, team processes, and organizational communication.