

Bridging the gap between learning and practicing leadership: Developing new instruments to
evaluate leadership and adaptation

By

Ahmadreza Shamsi Yousefi

B.A., University of Tehran, 2012
M.Sc., Iran's National University, 2015

AN ABSTRACT OF DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Staley School of Leadership

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Abstract

This dissertation aimed to narrow the gap between learning and exercising leadership. It tried to address three main gaps in the field of leadership development assessment as its main research objectives: 1) the measurement gap which concerns the divergence between academic and industry perspectives on leadership development 2) the teaching-assessment gap, involving the misalignment between the content taught and what is being assessed; and 3) learning and practicing environments gap which is about training-transfer. These gaps complicate assessing the true impact of leadership development. To bridge the gaps and in an engaged approach to research, the researcher chose the Kansas Leadership Center's (KLC) 2023 leadership development path as the research site for a mixed-method sequential design.

To address the first two gaps, the researcher engaged with KLC's faculty and staff to develop instruments grounded in adaptive leadership as the theoretical foundation of their teaching framework. Four main impact assessment tools were developed to track the reasoning processes leading to leadership intervention at the individual level: leadership and adaptation behavior (ALB) four-factor scale, adaptive leadership self-awareness (ALSA) three-factor scale, self-efficacy (ALSE) two-factor scale, and shifting-reflectiveness (ALSR) two-factor scale. The scales underwent rigorous validation steps, including content and face validity, exploratory and confirmatory factor analyses, criterion-related and construct validity, reliability checks, and generalizability tests. Drawing from Stanovich's (2011) tripartite model of reasoning, the researcher also developed a linear mixed effect model comprising of these scales that simulates the reasoning process underlying leadership and adaptation behavior. Moreover, in an evaluation study on KLC's 2023 path, this research presented a model finding that KLC exhibits balanced and unbalanced effects on the cognitive processes represented by these scales.

To tackle the third gap, the researcher engaged with KLC's participants through a maximum of two rounds of existential phenomenological interviews to delve into how individuals decide to exercise leadership in their professional settings. Building upon and corroborating certain findings from the quantitative phase regarding the cognitive processes entailed in leadership practice, the qualitative research elucidated the manner, context, and timing of these processes during the decision-making phase. The qualitative phase revealed that the reasoning process toward adaptive leadership intervention starts with cognitive dissonance and ends in a dialogic action space where individuals engage others to make multiple simulations about the problem at hand and design a collective solution. It also uncovered a post-intervention reflection/assessment phase that may offer an alternative reasoning path toward a more successful intervention.

Finally, the research provided specific recommendations for leadership development programs using adaptive leadership framework to consider bridging the gap between learning and practice. Emphasizing the conflict-ridden and emotion-laden nature of adaptive intervention, which may strain individuals' cognitive capacities and potentially lead them to withdraw from the intervention process, the researcher highlighted the significance of "relationality" in effectively navigating the intense moments.

Bridging the gap between learning and practicing leadership: Developing new instruments to
evaluate leadership and adaptation

by

Ahmadreza Shamsi Yousefi

B.A., University of Tehran, 2012
M.Sc., Iran's National University, 2015

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Staley School of Leadership

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Timothy Steffensmeier

Copyright

© Ahmadreza Shamsi Yousefi 2024.

Abstract

This dissertation aimed to narrow the gap between learning and exercising leadership. It tried to address three main gaps in the field of leadership development assessment as its main research objectives: 1) the measurement gap which concerns the divergence between academic and industry perspectives on leadership development 2) the teaching-assessment gap, involving the misalignment between the content taught and what is being assessed; and 3) learning and practicing environments gap which is about training-transfer. These gaps complicate assessing the true impact of leadership development. To bridge the gaps and in an engaged approach to research, the researcher chose the Kansas Leadership Center's (KLC) 2023 leadership development path as the research site for a mixed-method sequential design.

To address the first two gaps, the researcher engaged with KLC's faculty and staff to develop instruments grounded in adaptive leadership as the theoretical foundation of their teaching framework. Four main impact assessment tools were developed to track the reasoning processes leading to leadership intervention at the individual level: leadership and adaptation behavior (ALB) four-factor scale, adaptive leadership self-awareness (ALSA) three-factor scale, self-efficacy (ALSE) two-factor scale, and shifting-reflectiveness (ALSR) two-factor scale. The scales underwent rigorous validation steps, including content and face validity, exploratory and confirmatory factor analyses, criterion-related and construct validity, reliability checks, and generalizability tests. Drawing from Stanovich's (2011) tripartite model of reasoning, the researcher also developed a linear mixed effect model comprising of these scales that simulates the reasoning process underlying leadership and adaptation behavior. Moreover, in an evaluation study on KLC's 2023 path, this research presented a model finding that KLC exhibits balanced and unbalanced effects on the cognitive processes represented by these scales.

To tackle the third gap, the researcher engaged with KLC's participants through a maximum of two rounds of existential phenomenological interviews to delve into how individuals decide to exercise leadership in their professional settings. Building upon and corroborating certain findings from the quantitative phase regarding the cognitive processes entailed in leadership practice, the qualitative research elucidated the manner, context, and timing of these processes during the decision-making phase. The qualitative phase revealed that the reasoning process toward adaptive leadership intervention starts with cognitive dissonance and ends in a dialogic action space where individuals engage others to make multiple simulations about the problem at hand and design a collective solution. It also uncovered a post-intervention reflection/assessment phase that may offer an alternative reasoning path toward a more successful intervention.

Finally, the research provided specific recommendations for leadership development programs using adaptive leadership framework to consider bridging the gap between learning and practice. Emphasizing the conflict-ridden and emotion-laden nature of adaptive intervention, which may strain individuals' cognitive capacities and potentially lead them to withdraw from the intervention process, the researcher highlighted the significance of "relationality" in effectively navigating the intense moments.

Table of Contents

List of Figures	xi
List of Tables	xii
Acknowledgements	xiii
Dedication	xvi
Preface	xvii
Chapter 1 - Introduction	1
Strategic Research Site/Material (SRM)	2
Rationale	3
Research Objectives, Hypotheses, & Questions	5
Theoretical Commitments	7
Pragmatism and Constructivism	7
Overarching Theory	8
Content-Level Theories	9
Engaged Scholarship	9
Operationalization of Constructs	11
Subjectivity & Positionality	17
Limits and Possibilities of the Research	19
Chapter 2 - Literature Review & Theoretical Framework	21
Literature Review	21
Systematic Literature Review	21
Results of Co-Citation Analysis	21
Discussion of the Literature	28
Leadership Development Evaluation Clusters	28
Scale Development Clusters	31
Adaptive Leadership Cluster	32
From Heifetz to Uhl-Bien	32
Other Spots of the Literature	33
How the Literature Review Informed this Research	36
Theoretical Framework	40

Adaptive Leadership	41
Dual Process Theory of Reasoning (DPT).....	46
Tripartite Model of Mind	47
Connections Between Adaptive Leadership and DPT	50
Experiential Learning Theory (ELT)	51
ELT as the Connector of Adaptive Leadership and DPT	55
Theoretical Framework Components.....	57
Chapter 3 - Methodology	67
Hypotheses/Questions.....	68
Rationale	69
Methodological Framework.....	70
Research Design	77
Data Representation.....	96
Validity & Reliability of the Research	97
Engagement Strategy	101
Ethics	103
Chapter 4 - Findings & Discussion.....	105
Demographics	106
Quantitative Findings by Hypotheses	119
Qualitative Research Findings	169
Discussion of the Findings.....	190
Chapter 5 - Conclusion & Recommendations	216
Adaptive Leadership Measurement Tools	216
A Model for Measuring Impact	222
An Engaged Process of Theory Building	226
Recommendations.....	227
Recommendations for Adaptive Leadership Teachers	227
Recommendations for Engaged Scholars	232
Recommendations for Researchers	234
References.....	236
Appendix A - Item Generation	249

Appendix B - Final List of Items After Focus Group 1	253
Appendix C - Content Validity Results	256
Appendix D - EFA Results for ALB Scale	257
Appendix E - CFA Results for ALB Scale First Dataset (Model 1)	262
Appendix F - CFA Results for ALB Scale Dataset 1 (Model 2)	264
Appendix G - CFA Results for ALB Scale Dataset 2 (Model 1)	268
Appendix H - FA Results for ALB Scale Second Dataset (Model 2)	270
Appendix I - Convergent, Concurrent, Discriminant Validity	274
Appendix J - Focus Group Session Design	275
Appendix K - EFA Results for ALSE Scale	276
Appendix L - CFA Results for ALSE Scale First Dataset	279
Appendix M - CFA Results for ALSE Scale Second Dataset	282
Appendix N - Correlation Matrix (Convergent-Concurrent-Discriminant Validity)	285
Appendix O - EFA Results for ALSA Scale	287
Appendix P - CFA Results for ALSA Scale First Dataset	291
Appendix Q - CFA Results for ALSA Scale Second Dataset	294
Appendix R - EFA Results for ALSR Scale	297
Appendix S - CFA Results for ALSR Scale	300
Appendix T - Generalizability Test	303
Appendix U - Regression for Predictive Validity	305
Appendix V - LMM Models	315
Appendix W - YLE Impact on Competencies	321
Appendix X - LFC Impact on Competencies	334
Appendix Y - KLC's Path Model of Impact	347
Appendix Z - Research Protocols and IRB	355

List of Figures

Figure 2.1. Visual Map of Citation Clusters (Leadership development evaluation)	22
Figure 2.2. Visual Map of Citation Clusters (Leadership Scale).....	26
Figure 2.3. Visual Map of Citation Clusters (Adaptive Leadership).....	27
Figure 2.4. The Adaptive Leadership Process	41
Figure 2.5. The Lewinian Experiential Learning Model	51
Figure 2.6. From Elementary Modes to Integrative Complexity in Experiential Learning.....	54
Figure 2.7. Theoretical/Conceptual Framework	62
Figure 2.8. Adaptation and Complexity Stages	65
Figure 3.1. Scale Validation Process	80
Figure 4.1. Concept Map	173

List of Tables

Table 2.1. Cluster Information for Leadership Development Evaluation Documents	23
Table 2.2. Cluster Information for Leadership Scale Documents	25
Table 2.3. Cluster Information for Adaptive Leadership.....	27
Table 2.4. Theoretical Connections Between Adaptive Leadership, DPT, CDT, and ELT	61
Table 3.1. Criterion-Related and Construct Validity Instruments	83
Table 3.2. Pearson Correlation Coefficient Interpretation.....	83
Table 4.1. Demographic Characteristics of the Panel of Experts	106
Table 4.2. Demographic Characteristics of the EFA Dataset	107
Table 4.3. Demographic Characteristics of the CFA -Dataset 1.....	109
Table 4.4. Demographic Characteristics of the CFA- Dataset 2.....	110
Table 4.5. Demographic Characteristics of YLE Participants.....	112
Table 4.6. Demographic Characteristics of YLE Sample Participants	113
Table 4.7. Demographic Characteristics of LFC Participants	115
Table 4.8. Demographic Characteristics of LFC Sample Participants	116
Table 4.9. Focus Group 1 Participants.....	117
Table 4.10. Focus Group 2 Participants.....	117
Table 4.11. Demographic Characteristics of the Interviewees	118
Table 4.12. Component Tables and the ALB Items Kept.....	123
Table 4.13. ALB Factors Labels	129
Table 4.14. ALSE Components and the Items Kept.....	133
Table 4.15. ALSE Factors Labels	137
Table 4.16. ALSA Component and the Items Kept	140
Table 4.17. ALSA Factors Labels.....	145
Table 4.18. Component Tables and the ALSR Items Kept.....	148
Table 4.19. ALSA Factors Labels.....	152
Table 4.20. LMM Scenario Results	160
Table 4.21. LMMs Explanatory Strengths Change	160
Table 4.22. Code Book	170
Table 4.23. Qualitative Codes that Resonate with MRM and ALSR Items	209

Acknowledgements

I would like to extend my deepest acknowledgments to five groups of people whose unwavering support has been instrumental in fueling my research journey: the committee members, the faculty members in leadership communication program, the Kansas Leadership Center, my fellow leadership communication PhD cohorts, and last but certainly not least, my family.

This research journey would have been impossible without the guidance and support of my esteemed committee members: Drs. Tim Steffensmeier, Kerry Priest, Andy Wefald, and Becky Reichard. The content presented here is the culmination of countless hours of conversation and discussion, particularly with my advisor Dr. Steffensmeier, who has demonstrated an exceptional understanding of my work and consistently provided invaluable direction, motivation, and opportunities to advance this research. His patience with me in joining this program, adapting to the research work, and finishing this research is truly commendable.

Special gratitude is extended to Dr. Priest, whose humility, rich knowledge, and commitment to progress have served as a source of inspiration, both as a committee member and as the graduate program director.

I am profoundly grateful to Dr. Wefald for his role as a thought partner and motivator throughout this research. He has not only encouraged me to embrace the entire research process, with all its ups and downs, but has also shown me how to enjoy it.

Lastly, I am honored to have benefited from the thought provoking and challenging ideas of Dr. Reichard since the inception of this research. Her insightful comments, probing questions, and strategic suggestions at various stages of this dissertation have opened and continue to open new doors for exploration.

In addition to the knowledge and capacity-building opportunities provided by the leadership communication faculty and the Staley School of Leadership, this dissertation has directly benefited from the input of several faculty members during the courses I took with them. From Dr. Tim Shaffer who instilled in me the joy of delving deep into literature to Dr. Brandon Kliewer who provided invaluable insights in solidifying this study's theoretical framework. I am also grateful to Drs. Alisa Garni and Gary Brase for their significant contributions to the research design and methodologies employed in this study. Special thanks to Dr. Craig Paul for sparking my interest in mixed-method research, and to Drs. Colene Lind and David Procter for offering opportunities to acquire the knowledge and practical skills necessary for designing dialogic and deliberative spaces within this research. Lastly, I am indebted to Dr. Susan Yelich Biniecki for her guidance in framing and writing up the qualitative results. Their collective input has greatly enriched this dissertation.

This research was made possible through the meaningful collaboration with Kansas Leadership Center, serving as our research partner. I am deeply appreciative of the active engagement of their teachers, staff, and participants throughout the research process, from data collection to interpreting the results. Special thanks are also extended to Third Floor Research for providing the institutional platform that facilitated the collaboration between Staley School of Leadership and KLC throughout this research.

I am immensely grateful for the constant assistance and energy I received from all the Leadership Communication PhD cohorts. From the cohorts of 2018-2019, whom I had eagerly counted the days to join, to the cohorts of 2020-2022, with whom I shared the learning space and experienced the joy of learning together. I cannot overstate the support I received from my friends, classmates, and colleagues, including Roberta Maldonado Franzen, RJ Youngblood,

Saya Kakim, Mac Benavides, Tamas Kowalik, Ania Payne, Innocent Assoman, Shukeyla Harrison, Chibuzor Azubuike, Josh Brewer, Cale Morrow, Anisah Ari, Phil Black, Trysta Morenz, Carlie Snethen, and Emmanuel Jeje. Whether during moments of celebration and success or during challenging times of loss and failure, their presence and support have been invaluable.

I extend unlimited thanks to all those who generously offered mentorship, coaching, and valuable insights throughout this research journey. I would like to particularly highlight the contributions of Mac Benavides and Jeff Ramdass, who dedicated a significant amount of their time to provide teaching and coaching, aiding me in navigating the complexities of qualitative and quantitative analyses. Additionally, thanks to all the 21 members of the panel of experts who contributed their time providing insights into the foundational aspects of this research.

Finally, I would like to express my heartfelt appreciation to my wife, Ellie Salehi, for selflessly sacrificing her time and energy to support me throughout this journey and my family members back in Iran who have never stopped their encouragement, despite the miles that separate us. Their genuine love and belief in me have been the pillars of strength and hope throughout this endeavor.

Dedication

To my mom who sowed the seeds of passion for learning within me, but departed when I just started this research.

Preface

The initial flames of interest in this research ignited before coming to the United States when I was working in Iranian slums and informal settlements as a leadership coach. While training civic practitioners, I often encountered feedback suggesting that the content presented in workshops, books, and articles was challenging to apply in practice and did not necessarily align with real-world realities. Practitioners would express sentiments such as, “The modules and content you provide in the workshops are insightful but difficult to implement in the field.” Some even became discouraged, feeling they were not making a meaningful impact in addressing public challenges. Departing Iran, I carried with me the question: “Did I really impact the practitioners and add to their learning?”

Moving to the US and working on the impact assessment of Kanas Leadership Center’s leadership development programs through Third Floor Research deepened this question in my mind. Conversations with many leadership development educators and immersion in people’s stories of leadership practices reinforced a prevalent phenomenon: while our teachings may initially spark enthusiasm for learning, challenges inevitably arise when attempting to apply these lessons. This observation persisted as a concern throughout my collaboration with Third Floor Research and ultimately shaped the focus of this dissertation.

The primary audience for this research includes leadership trainers and educators who aim to bridge the gap between their kind learning environments and wicked practice environments. Additionally, it might receive special attention from leadership development programs utilizing adaptive leadership framework, seeking to assess their effectiveness and enhance impact. Furthermore, scholars investigating the theoretical foundations of adaptive leadership and working to prove it as a theory may find this research compelling. Finally, the

engaged scholars focused on program evaluation might also find this research relevant as it conducted a collaborative process in developing new assessment tools and adopted a formative approach to evaluation in partnership with Kansas Leadership Center.

Keyhan Shams

Manhattan, Kansas

Nowruz 13, 2583 | April 1, 2024

Chapter 1 - Introduction

Interest in leadership development programs (LDPs) has increased over the last two decades. O'Leonard and Krider (2014) reported a 14% increase in leadership development investment in 2014 compared to previous years by the US companies. *Chief Learning Officer* magazine stated that the increase continued up to 2017 among corporations. Moreover, according to Training Industry (2022), organizations' investment in leadership reached above 370 billion dollars (about \$1,100 per person in the US) worldwide in 2019. However, the effectiveness of LDPs has usually been a concern among funders. According to Leadership IQ study (2020), typical leadership development programs have not been motivating for many companies to improve their employees' potentials. This might suggest that there is still a sense of reluctance in the eyes of a group of organizations to continue such development programs.

Program evaluation studies can help companies and funders in addressing their hesitation about LDPs. However, these studies have faced their own challenges:

- 1) First, while evaluation studies may show an increase in individuals' leadership competencies right after the program, this might not be observed in the real world. A phenomenon that is called the training-transfer gap. One possible reason behind this might support the idea that learning settings in LDPs do not perfectly match the target settings where people make predictions, choice, and act (Hogarth et al., 2015).

Besides any possible individual reasons behind this mismatch, the dynamic actual environments may not represent what people learn in training programs. We call this here as the *learning and practicing environments gap*.
- 2) Second, there is a gap between the ways academics and industries assess LDPs.

Crawford and Kelder (2019) declared that there is a clear divide between academic

assessment and industry tools/scales in the field of leadership development. We call this here the *measurement gap*.

- 3) Third, there might be a gap between what is taught in the LDPs and what is assessed. We call it here the *teaching-assessment gap*. This relates to the validity and reliability of evaluation studies and questioning whether the measurement tools assess what is being taught in the programs (Newcomer et al., 2015). And if so, does it reproduce the same results for other similar programs?

As a scholar-practitioner working in communities, companies, and academia over the last 10 years, the researcher faced all these challenges during his professional and educational life. Working with people and change agents in Iranian slums and informal settlements as a leadership coach, he heard people frequently telling him “What you teach does not match with our real-world experiences” (*Environment gap*). As a scholar and consultant for the Iranian government, the researcher received feedback from officials expressing that the impact measurements he would employ did not align with their assessment criteria (*Measurement gap*). Likewise, as a graduate research assistant tasked with evaluating the Kansas Leadership Center’s (KLC) LDPs, he has been involved in evaluation studies in which evaluation metrics did not accurately reflect the teaching modules (*Teaching-assessment gap*). Therefore, these three challenges came as an important phenomenon to the researcher to explore and address in this research.

Strategic Research Site/Material (SRM)

To address these challenges, the researcher chose KLC’s 2023 leadership development path as the strategic research site (i.e., “an empirical material/site that exhibits the phenomenon in such an accessible form that it enables fruitful investigation of problem for further inquiry”,

Merton, 1987, pp.11-12). KLC's 2023 path had three steps consisting of three main programs: Your Leadership Edge (YLE), Lead for Change (LFC), and Equip to Lead. YLE and LFC were selected as the focus of this research to address the aforementioned gaps for three main reasons:

- 1) The researcher has been working with KLC since 2021, and he is familiar with their teaching methodology and adaptive leadership as a framework informing their teaching.
- 2) Currently, there are no validated assessment tools developed for measuring adaptive leadership competencies except Raei's (2018) scale which is more focused on the authority aspects of adaptive leadership.
- 3) KLC's past evaluation studies were not completely aligned with its teaching framework. Before 2021, KLC's impact assessment was primarily limited to its learning objectives. In 2021, KLC's impact assessment took a research-based approach (Shams et al., 2023). However, the assessment tools (transformational leader behavior and self-views) did not fully align with KLC's leadership framework.
- 4) KLC's path program was designed with this aim to reduce the training transfer gap. At the end of the path, participants are equipped to intervene in their work settings.

As such, this SRM, involves intrinsic and instrumental interests (Merton, 1987) for the researcher. Intrinsic interests stem from the alignment of concepts being used in the path with the researcher's personal interests, especially adaptation and adaptive leadership. In terms of instrumental interests, KLC is facing the program evaluation challenges discussed above.

Rationale

These challenges have also been highlighted in leadership literature. Leadership scholars have been working to find ways to bridge these gaps. For instance, regarding the environment gap, Leadership-as-Practice theory (LAP) contends that most actions happen in the spot, in situ,

and on feet. Therefore, it proposes that leadership practice resides in people's unconsciousness or dwelling mode of engagement as opposed to consciousness or building mode (Carroll et al., 2008; Heidegger, 1962). As such, LAP suggests merging the two settings by making leadership development in situ, implicit, and unconscious (Chia & MacKay, 2007; Denyer & James, 2016; Raelin, 2011). In contrast, inspired by Lewin's learning model, the adaptive approach to leadership emphasizes the role of the conscious mind and detached reflection space in learning and practicing leadership. It explains that these target environments contain adaptive challenges that require case-specific learning, taking risks, and accepting change in priorities, beliefs, and habits that people typically do not practice and generate work avoidance (Heifetz, 1994; Heifetz & Linsky, 2002; Heifetz et al., 2009). Hence, teaching approaches like *case-in-point* needed to bring this consciousness (heat) to people that the existing plan is not enough for new situations.

Regarding the measurement gap, Newcomer et al. (2015) suggested involving stakeholders in evaluation design. According to Newcomer et al. (2015), "as professional evaluation standards prescribe, program stakeholders should also be involved to ensure that the results of evaluation work of any kind will be used" (p. 11). Finally, regarding the validity and reliability of the evaluation scale, Crawford and Kelder (2019) offered 18 critical checks for four stages of leadership scale development: theory generation, item development, content validity, and empirical evaluation.

Although these recommendations have been discussed in leadership, an actual program evaluation conducted to address these three challenges can hardly be found. This research tried to address these three challenges in an engaged program evaluation focusing on KLC's 2023 leadership development path as the strategic research site. The methodological approach of this study could be used by program evaluation designers regardless of the field. The theoretical and

empirical findings of this research could be employed by leadership researchers, educators, and practitioners.

Research Objectives, Hypotheses, & Questions

This research followed a consistent hierarchy of goal, aim, objectives, hypotheses, and question to address the three main challenges discussed above regarding LDP assessment. The goal of this research was to narrow down the gap between learning and practicing leadership. The aim was to develop a valid, reliable, generalizable, and practice-reflective assessment model for adaptive leadership. Following these goal and aim, specific and pertinent objectives, hypotheses, and questions have been designed.

Objective 1: Developing Instruments for Evaluating LDPs with Adaptive Approach

This objective was defined to address the *teaching-assessment* and *measurement gaps* and has one main hypothesis comprising four subsections.

Hypothesis 1: Adaptation and leadership or adaptive leadership assessment scales below demonstrate acceptable levels of content validity (CV) (face validity, simplicity, relevancy, clarity, necessity), construct validity (convergence and discriminant validity as well as exploratory factor analysis (EFA) and confirmatory factor analysis (CFA)) (criterion-related validity (concurrence and predictive), and internal consistency/reliability tests (Cronbach test).

- 1.1) Self-efficacy (ALSE)
- 1.2) Shifting reflectiveness (ALSR)
- 1.3) Self-awareness (ALSA)
- 1.4) Leadership behavior (ALB)

Objective 2: Measuring the Effectiveness of KLC's Leadership Development Path

Applying a longitudinal approach, this objective was defined to address the *teaching and practicing environments gap* and has one main hypothesis.

Hypothesis 2: Participation in KLC's path leads to significant positive change in participants' adaptation and leadership competencies (ALB, ALSA, ALSR, ALSE, and master rationality motive (MRM))

Objective 3: Providing Specific Recommendations for Teaching

This objective was adopted to address the teaching and practicing environment gap. It consists of two quantitative and qualitative phases. The quantitative phase reveals what elements contribute significantly to exercising leadership. Moreover, it provides the teaching team with concrete feedback from the practicing environment through the qualitative phase of the research. Additionally, the qualitative findings were used to explain the quantitative results and provide more details about the involvement of reflective mind components in adaptive leadership behavior.

Hypothesis 3: There is a positive relationship between adaptive leadership behavior (ALB) and Stanovich's reflective mind components (MRM, ALSR, ALSE, and ALSA): (Regression model: $Y_{ALB} = \beta_0 + \beta_1 X_{SE} + \beta_2 X_{MRM} + \beta_3 X_{SR} + \beta_4 X_{SA} + \epsilon$)

Qualitative Research Question: 1) How do KLC participants describe the decision-making process they take to exercise/not to exercise leadership?

Theoretical Commitments

The ontological and epistemological commitments of this research are introduced in this section. Chapter 2 explains the theoretical framework in more detail.

Pragmatism and Constructivism

This study is located at the interplay space between positivism and interpretivism (Ospina & Uhl-Bien, 2012). This mixed-method research and its theoretical framework is situated at the intersection of pragmatism and constructivism. In pragmatism, especially in Pierce's view, there is a reality independent of the mind. However, it is accepted as far as it guides the action. In pragmatism, the process of theorizing is important and unlike logical positivism, it does not start with induction, but intuition. In pragmatism, experience is the unit of study.

This research followed this pragmatic tradition. In the quantitative phase, particularly, there are two main hypotheses (1 & 3) informed by the content level theories/models (Dual process theory of reasoning and adaptive leadership) that were tested in the KLC's 2023 path evaluation study. This hypothesis testing did not happen through objective fact-finding but by focusing on the individual's self-report of their leadership experiments. For instance, in testing hypothesis 1, although the adaptive leadership framework might have different dimensions, only those that passed the validation statistical tests (based on people's responses) were kept as assessment factors. From a pragmatic perspective, those discerning items that guide assessment remained in the measurement tool.

In the sequential qualitative phase, the researcher conducted one-on-one existential phenomenological interviews without any theory in mind and focused on the participants' stories of leadership experiments (Van Manen, 2014). This approach added a constructivist dimension to this research as people reported a mental construction of their leadership activities in the

interviews (Schwandt, 1994). Particular focus was given to concrete stories of leadership interventions without asking for any analytic reflection. In doing so, the researcher looked for evidence that supports both the results of the quantitative phase and/or explains the surprises. Both this evidence took the study deeper into the reasoning processes involved in leadership behavior and provided insights for adaptive leadership teachers.

In sum, this research did not solely take a verification process, but also a falsification and abduction process mixing the two phases. In this view, there is a reflective relationship between the informing theory and practice. Therefore, the theory takes a Darwinistic evolutionary process to grow. Through this process, it would be revised and updated based on the surprising research results and findings of both quantitative and qualitative phases (Van de Ven, 2007).

Overarching Theory

This research is informed by constructive developmental theory (CDT) as the overarching umbrella. Drawing upon Piaget's (1954) developmental growth, CDT was initially proposed by Kegan (1982). However, it took different forms alongside and later especially in the works of Kohlberg (1969) and Torbert (2004). This research employed Kegan's (1982, 1994) CDT and its latest conceptualization regarding cognitive complexity, adaptation, and three orders of mind: socialized mind, self-authoring mind, and transformative mind as outlined by Kegan and Lahey (2009). The primary reason for incorporating this theory into the theoretical framework was Kegan's (1994, 2009) emphasis on cognitive complexity and his argument about the necessity of a self-transforming mind in adaptive leadership. Therefore, Kegan's proposition that adaptive leadership necessitates a complex mindset provides a framework for examining whether the KLC's program enhances the cognitive complexity of its participants.

Content-Level Theories

This research benefits from three content-level theories: Adaptive Leadership, Experiential Learning Theory (ELT), and Dual Process Theory (DPT). Adaptive leadership responds to this “what” question: What is adaptive leadership behavior? KLC is trying to build adaptive capacity in people. Therefore, to measure KLC’s impact at the individual level, it needs to be defined, first. This can help us with developing measurement tools and scales. DPT responds to this “how” question: How is it performed individually? What is the mechanism behind it? This can be explored from different perspectives. In this study, we try to address it from a cognitive psychology perspective. Responding to this “how” question can help us inform KLC about how to tailor their programs to the people’s reasoning process involved. ELT, here, is being used as a connector theory that responds to this “Why” question: Why do we think DPT explains the processes involved in adaptive leadership behavior?

In other words, this study incorporates three theories of mind, meaning, and truth based on Meyers’ (1999) pragmatism. DPT is the theory of mind that posits the hypotheses and ideas of rational action. ELT is the theory of meaning that reveals the connection between the theory of mind and action. Finally, adaptive leadership is the theory of truth which shows whether the hypotheses are true in guiding action and prediction. Any surprise arising from the theory of action necessitates modification in the theory of mind.

Engaged Scholarship

This research took an engaged *summative to formative* approach to program evaluation. According to Newcomer et al. (2015), “Formative evaluation uses evaluation methods to improve the way a program is delivered...[while]...summative evaluation measures program outcomes and impacts during ongoing operations or after program completion” (p. 11). In this

research, program clients and providers were involved in the evaluation process. KLC's teaching team and participants were involved in the development of measurement tools. KLC's teaching team was invited to focus group discussions about the measurement tools' CV, refining, and making sense of scale constructs/factors. Moreover, KLC's path participants were involved in the qualitative phase of the research to inform us about how KLC can benefit from the quantitative evaluation results. Their concrete stories helped the researcher to explain the results of the quantitative phase and provide more details about the assessment tools and models. Moreover, their contribution assisted in providing specific recommendations to KLC's teaching team and connecting the learning and practicing environments.

From Van de Ven's (2007) view of engaged scholarship, this research falls under design and evaluation form of engaged scholarship. Stakeholder engagement in this research, particularly KLC, ensured their consent and influence over the measurement tools, thereby fostered their commitment to utilizing the findings for enhancing their LDPs. Different steps of this research illustrated three main components of Van de Ven's (2009) diamond model: problem formulation (Based on KLC's needs), theory building (sense making in collaboration with KLC and their participants), problem solving (KLC's teaching team).

Operationalization of Constructs

Here, the concepts, constructs, and variables used in this research and its theoretical and methodological framework were defined.

Adaptation/Adaptive Work

For adaptive work, we use the definitions from Ronald Heifetz and Keith Stanovich. Heifetz (1994) defined leadership as “activity to mobilize adaptation” (p. 27), and by adaptation he goes beyond biological level which is about coping and survival without planning. According to Heifetz (1994),

In human societies, adaptive work consists of efforts to close the gap between reality and a host of values not restricted to survival. We perceive problems whenever circumstances do not conform to the way we think things ought to be. Thus, adaptive work involves not only the assessment of reality but also the clarification of values. (p. 31)

Heifetz (1994) defined adaptation as more active and deliberative rather than passive and merely coping with the situation. In his view, adaptation in an organizational context is a process of conservation and loss that takes time, relies on diversity, and occurs in experiments (Heifetz et al., 2009). Another aspect of adaptation is highlighted by Heifetz et al. (2009) when they say adaptive leadership is about “skill plus will” (p. 37). Later, it will be discussed how this definition connected adaptive leadership to other theories in the theoretical framework.

Heifetz’s definition of adaptation speaks to Stanovich’s (2011) definition of adaptation. He presents a critical perspective to the dominant definition of adaptation, “The tendency to make judicious decisions that serve one’s goals” (p. 175). Stanovich (2011) asserted that this definition limits adaptation to intelligence and instrumental rationality. Covering adaptation in the definition of rationality, through empirical research, Stanovich (2011) defined

adaptation/rationality “as the actions of an entity in its environment that serve its goals. It is indexed by thinking dispositions tapping individual differences in the reflective mind as well as fluid intelligence, which taps individual differences in the algorithmic mind” (p. 176). The concepts of algorithmic and reflective minds are separately discussed later in this chapter.

Heifetz’s focus on the combination of human values as well as “skill plus will” in his definition of adaptation and adaptive leadership aligns with Stanovich’s emphasis on thinking dispositions, and “algorithmic mind plus reflective mind” in his definition of adaptation/rationality. This perspective toward adaptation is used in this research to conceive of ALB as rational action.

Adaptive Challenge

Distinguishing from technical challenges, Heifetz introduced various aspects of adaptive challenges in his 1994, 2001, and 2009 books. Among all, these two verbatims provide insights for this research.

- 1) A particular kind of problem where the gap between reality and values/aspirations cannot be closed by the application of current technical know-how or routine behavior. To make progress, not only must invention and action change circumstances to align reality with values, but the values themselves may also have to change. (Heifetz, 1994, p. 35)
- 2) [Those challenges that] “can only be addressed through changes in people’s priorities, beliefs, habits, and loyalties. Making progress requires going beyond any authoritative expertise to mobilize discovery, shedding certain entrenched ways, tolerating losses, and generating the new capacity to thrive anew” (Heifetz et al., p. 19).

The concept of adaptive challenge has been reiterated under various terminologies in the literature. For instance, Rittel and Webber's (1973) definition of tame and wicked problems is similar to Heifetz's technical and adaptive challenges. However, Raei (2018) argued that there are a few differences in terms of field context and essence of challenges. This discussion does not fit the scope of this research.

Cognitive Dissonance

Festinger's (1957) definition of cognitive dissonance is used in this research. According to Festinger (1957), dissonance and consonance describe the relational dynamics between pairs of elements. He defines elements here as cognition and defines cognition itself as "The things a person knows about himself, about his behavior, and about his surroundings" (p. 9). In other words, he defines elements as knowledge. Festinger (1957) took knowledge as well as values, opinions, beliefs, and attitudes as elements of cognition. Then, he says "Two elements are dissonant if, for one reason or another, they do not fit together. They may be inconsistent or contradictory, culture or group standards may dictate they do not fit, and so on" (pp. 12-13). In this research, we focused on these inconsistent elements in the work environments as part of the reasoning process for exercising new leadership behavior.

Type 1 Mind: The Autonomous Sets of Systems (TASS)

TASS or Type 1 (Stanovich, 2008, 2011) or System 1 (Kahneman, 2011) is automatic, peripheral, impulsive, and a fast level of mind that relies on well-learned information and has roots in human's routine activities, habits, and recurring tasks of daily life. In contrast, Type 2 or System 2 is slower, more deliberative, analytic, cautious, lazy, and slow and relies on rules and symbolic logic (Evans, 2008; Schneider & Shiffrin, 1977; Stanovich, 1999, 2008; Kahneman,

2011; De Neys, 2018). Stanovich (2008, 2011) used the word TASS per his argument that this is not a singular automatic system but plural.

Type 2 Mind: Algorithmic & Reflective Mind

Type 2 level of mind was defined in the previous section. Stanovich (2008) brought more complications to DPT by diving deeper into this type of mind and introducing *algorithmic* and *reflective* levels of mind:

TASS will implement its short-leashed goals unless overridden by the algorithmic mechanisms implementing the long-leash goals of the analytic system. But override itself is initiated by higher level control. That is, the algorithmic level of the analytic system is conceptualized as subordinate to the higher-level goal states and epistemic thinking dispositions [that] ...exist at what might be termed the reflective level of processing—a level containing control states that regulate behavior at a high level of generality. (p. 57)

To put it differently and tying it back to the concept of adaptation, from Stanovich's perspective, adaptation involves a dynamic between two distinct cognitive processes within reflective and algorithmic minds in which the former plays the role of commander for the latter as the operator.

Cognitive Decoupling

Leslie (1987) defined decoupling as modeling a secondary representation. He called it representation of representations or *metarepresentation* of the world. Informed by Leslie's (1987) conceptualization, Stanovich (2008) defined cognitive decoupling as "copy[ing] the primary representation ...that was decoupled from the world so that it could be manipulated" (p. 62). This research explains more about how decoupling happens during the reasoning process toward leadership behavior.

Master Rationality Motive (MRM)

According to Stanovich (2011), “MRM is the motive that drives the search for rational integration across our preference hierarchies” (p. 81). Stanovich (2011) introduced MRM as a scale to measure reflective mind and thinking dispositions. In other words, MRM is the motivation for cognitive decoupling particularly cognitive decoupling 1 & 2 (inhibition and updating).

Shifting Reflectiveness

According to Stanovich (2011), the third type of decoupling which sustains making multiple simulations and integrates them toward action is called shifting as the executive functions. It is commanded by thinking dispositions that here we relate them to Day et al.’s (2009) goal orientations. As part of the reflective mind processes, here we called it adaptation and leadership shifting reflectiveness (ALSR).

Leadership Self-Efficacy

Perceived self-efficacy was defined for the first time by Bandura (1986) as “people’s judgments of their capabilities to organize and execute courses of action required to attain designated types of performances” (p. 391). Bandura (2006) suggested this construct “must be tailored to the particular domain of functioning that is the object of interest” (p. 308). As such, we use Hannah et al.’s (2012) definition of leader self-efficacy that is “leaders’ level of perceived capability to self-regulate their thoughts and motivation, draw from means in their environment, and act successfully across a span of leader challenges and tasks in their current context” (p. 1). This definition is important as it puts self-efficacy under reflective mind processes. In fact, here, we focus on action component of self-efficacy that is interpersonal and the self-regulatory self-efficacy that contains cognitive processes internal to the leader (Hannah

et al., 2012): “efficacy to plan and structure activities, to enlist needed resources, to manage the emotionally and cognitively disruptive effects of obstacles, setback, and stressors” (Bandura, 1997, p. 53). Hannah et al. (2012) argued that self-efficacy should go beyond perceived capability of leaders to act and incorporate their capability to interpret the complex problems of the context and create solutions. Here, we labeled the adaptation and leadership self-efficacy as ALSE.

Apprehension

According to Kolb’s (1975, 1984, 2014) definition, apprehension is an adaptive mode or style of learning that happens through concrete experience (CE) of our immediate environment. According to Kolb and Fry (1975), “concreteness represents the absence of [abstractness] – the immersion in and domination by one’s immediate experiences” (p. 36). In other words, Kolb’s concrete CE means putting self openly in new experiences without biases and predispositions.

Comprehension

According to Kolb’s (1975, 1984, 2014) definition, comprehension is an adaptive mode or style of learning that happens through abstract conceptualization (AC) and grasping the experience. Kolb (1975, 1984) explained that, in AC, people turn their observations and CE into sound theories, conceptual interpretations, and symbolic representations to be used in their decision making.

Intention

Kolb (1975, 1984, 2014) defined intention is as an adaptive mode or style of learning that happens through reflective observation (RO). According to Kolb (1984), intention is one way to “transform the figurative representation of experience through internal reflection” (p. 41). In fact,

through RO, people try to reflect on their concrete experiences from many perspectives (Kolb & Fry, 1975).

Extension

Kolb (1975, 1984, 2014) defined extension as an adaptive mode or style of learning that happens through active experimentation (AE). Kolb and Fry (1975) stated that, in this mode, people “actively test the implications of their hypotheses” (p. 37). According to Kolb (1984), besides intention, extension is another way to transform the symbolic representations through active external manipulations in the real world.

Subjectivity & Positionality

The researcher’s subjectivity and positionality might have affected the qualitative phase of this research. First, his personal interests might have contaminated the research. Based on his research experiences in the past, it is likely that the analysis and interpretations were influenced by his recent studies in cognitive psychology. This is referred to as *availability bias*. Since he has such a proclivity for using the latest interesting concepts and ideas read or heard in the texts and speeches, it is probable that these concepts have caused myopia.

Second, the researcher’s previous work experience and living environment might have affected the research. Because of his background in urban planning and working in marginalized areas, it is possible that his research sometimes falls into fallacy and loses its neutrality, especially during interviews when people talked about their position within the power hierarchy.

Third, the researcher’s positionality as a graduate research assistant for Third Floor Research, a partnership between the KLC and Kansas State University, might have affected both data collection and analysis. He might have looked for positive signs of KLC’s effectiveness and

people might have just mentioned the positive side of their experiences to the KLC interviewer as the representative of their organization's partner.

Fourth, while doing sequential explanatory design, there is a potential risk of the researcher falling into confirmation bias. In this case, the researcher might have interpreted qualitative results in a way that confirms the quantitative results.

Subjectivity is a garment that may never be removed (Crotty, 1998). While the researcher cannot fully eliminate these subjectivities and biases from this research, being critically aware of them helped him protect the research from being too contaminated. Below are the three strategies that the researcher took to control his subjectivity and positionality in this research project.

First, in his interviews he insisted that any feedback from the interviewees is valuable for the research and improving the KLC's future programs. Doing so, he tried to reduce the possibility of making fake success stories by the participants who tended to show conformity.

Second, he tried to challenge the findings democratically (Timmermans & Tavory, 2012). Interviewees were provided with transcriptions and analysis for final verification and member check.

Finally, the researcher tried to cultivate the expectation within himself to always seek out surprises. Knowing that the quantitative phase and the content-level theories may direct the qualitative analysis and being aware of his tendency to associate concepts as a thinking error, the researcher constantly tried to find the findings that disrupt the quantitative results. He also convened several sessions with his major advisor and committee members to discuss findings that might have emerged out of his subjectivity/positionality.

Limits and Possibilities of the Research

This research has certain limitations regarding the researcher's subjectivities, generalizability, and data collection. The subjectivity limits were discussed elaborately in the previous section. In the following paragraphs, other limits are discussed.

Generalizability in quantitative research shows whether the research results, in this research tools/measures/scales, are applicable in other contexts, populations, and situations. It is important since it adds to the external validity of a research or instrument. Cronbach et al.'s (1963) generalization theory considers person, activity, items, time, and setting as sources of variation or error that researchers may consider making their results generalizable. Since this research was conducted only in Kansas and the scale items were developed in English, it is difficult to say that these measures would work in all contexts or for all LDPs. However, there is opportunity for future studies to replicate the research and scales in other LDPs, languages, and contexts.

Collecting enough longitudinal data was a major concern especially in the quantitative phase. Funded by Third Floor Research, at the later stages of the research, we decided to offer incentives for the T4 (four months after YLE) survey. Moreover, collecting data through self-report surveys may also increase the possibility of common method bias (Podsakoff, et al. 2003). To address this issue, the researcher followed Podsakoff et al. (2003) recommendation by asking participants at multiple time points. Moreover, for the qualitative research, two courses of interviews were scheduled with participants at two different times. However, due to time limits and especial circumstances, in some interviews, the two rounds were conducted at the same time.

Highlights of the Chapter

- Three major gaps in program evaluation were introduced: 1) Measurement gap; 2) learning and practice environments gap 3) Teaching-measurement gap.
- Goal, aim, and objectives, hypotheses, and research question were outlined in relation to the research gaps.
- The reason why KLC's 2023 path was used as the SRM was explained.
- Important and key constructs were operationalized.
- The researcher's positionality as the KLC's research assistant as well as his cognitive biases as subjectivities were highlighted.
- Limitations of research were discussed.

Chapter 2 - Literature Review & Theoretical Framework

Literature Review

Systematic Literature Review

This research utilized document co-citation analysis as a bibliometric method to conduct a systematic literature review. Document co-citation analysis investigates the origins of a specific field by identifying groups of authors who are frequently cited together. It measures the frequency with which two documents within the field are co-cited. It indicates the strength of the co-citation links and sheds light on the prevailing discourse in that field (Vogel, 2012; Vogel et al., 2021).

This research sought to develop new instruments to evaluate the impact of KLC's leadership development path, rooted in adaptive leadership framework. Hence, the bibliometric analysis is primarily focused on three main topics: "leadership development evaluation", "leadership evaluation scales", and "adaptive leadership". Employing document co-citation analysis as the primary method, the research aimed to explore the clusters of documents associated with these three focal areas. It is important to trace these areas in the literature to identify the theoretical gap especially in adaptive leadership and inform research methodology particularly in the field of program evaluation.

Results of Co-Citation Analysis

Leadership Development Evaluation

Data Collection from Web of Science. For the field of leadership development evaluation, the key words of "Leadership development" and "Evaluation" were searched. The search resulted in 2932 documents consisting of all types of documents (articles, book chapters, review papers, etc.) in all default Meso and Micro topics. Then, reviewing the abstracts of these

articles, the researcher ran a refinement process that resulted in 476 documents. The documents that were not related to leadership or were not focused on program evaluation were removed from the list. The data was exported into plain text and CSV file to be visualized in the next step.

Visualization. To facilitate the cluster and strength analyses of the collected data from Web of Science, the researcher used VOSviewer software. This software allowed for the visualization of the data. The analysis focused on co-citation as the type of analysis, with cited references serving as the unit of analysis. The threshold for inclusion was set at a minimum of five citations.

The visualization generated 171 main documents that formed five clusters. After conducting an initial review of the cluster documents, the list was refined to the 100 most relevant documents, which were organized into three main clusters (see Figure 2.1 and Table 2.1). Cluster 3 (blue), which was found to be more closely related to the field of health leadership, was removed from the review list. The remaining two main clusters were retained for further examination. Table 2.1 provides a brief overview of the most important documents, their main themes and link strength values.

Figure 2.1. Visual Map of Citation Networks and Clusters of 100 documents (Leadership development evaluation)

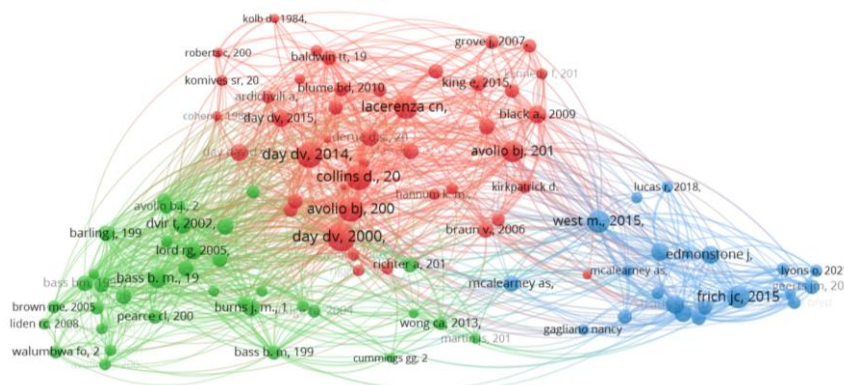


Table 2.1. Cluster Information for Leadership Development Evaluation Documents

Clusters	Authors	Strength	Main themes
1) Paradigm shifts	(Day, 2000)	192	Difference between leader development and leadership development Different practices of leadership development and necessity of assessment The necessity of connection between human capital and social capital
	(Day et al., 2014)	159	Process and content issues in leadership development-evaluation methods in leadership development longitudinal research in leadership development
	(Collins & Holton, 2004)	158	The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001
	(Avolio et al., 2009)	142	Leadership intervention definition Types of leadership intervention outcomes: Affective, cognitive, behavioral, organizational performance Transformational leadership theory vs traditional theories' impact on outcome variables Alignment of dependent variables and intervention theory in evaluation
	(Lacerenza et al., 2017)	139	How to design an impactful leadership training Alignment of outcomes with assessment criteria
	(Avolio et al., 2010)	109	Return of leadership development investment (RODI)
	(Day & Dragoni, 2015)	77	Human development theories: Dynamic skill theory and CDT self-views as proximal variables for predicting distal outcome variables Relational perspectives in leadership development: Transformational leadership, shared/distributed leadership Leadership as a system of adaptation
2) Measurement approaches	(Bass, 1985)	90	Emotional components of transformational leadership Differences between transformational leadership and transactional leadership
	(Dvir et al., 2002)	85	Longitudinal evaluation using randomized field experiment Direct vs indirect leadership Follower development
	(Burns, 1978)	76	Transformational leadership theory explained Individualized consideration Intellectual stimulation Inspirational motivation Idealized influence
	(Black & Ernest, 2009)	76	Designing an evaluation framework and scale validation
	(Lord & Hall, 2005)	73	Identity, meta-cognitive processes, and emotional regulation as main factors in developing the deeper cognitive structures involved in leadership expertise
	(Bass & Avolio, 1990)	58	Training transformational leaders Basic/ advanced workshops Evaluating workshops
	(Avolio & Hannah, 2008)	48	Leader developmental readiness constructs: learning goal orientation, developmental efficacy, self-concept clarity, self-complexity, metacognitive ability Training programs as trigger events adaptive vs maladaptive self-reflection

Leadership Scale

Data Collection from Web of Science. A similar process was completed for leadership evaluation scale area. The key words of “leadership scale” was searched in Web of Science which resulted in 7774. The researcher ran two rounds of refinement. In the first round, the results were filtered based on type of document (Article, review article, book chapter, paper proceedings) and relevant fields related to leadership (management, psychology, educational leadership, business, etc.). This refinement generated a list of 3319. Then, in the second round of refinement, the researcher went through the abstracts of these documents and omitted documents irrelevant to leadership evaluation scale. This gave birth to the final list of 1000 documents. The data was exported into plain text and CSV file to be visualized in the next step.

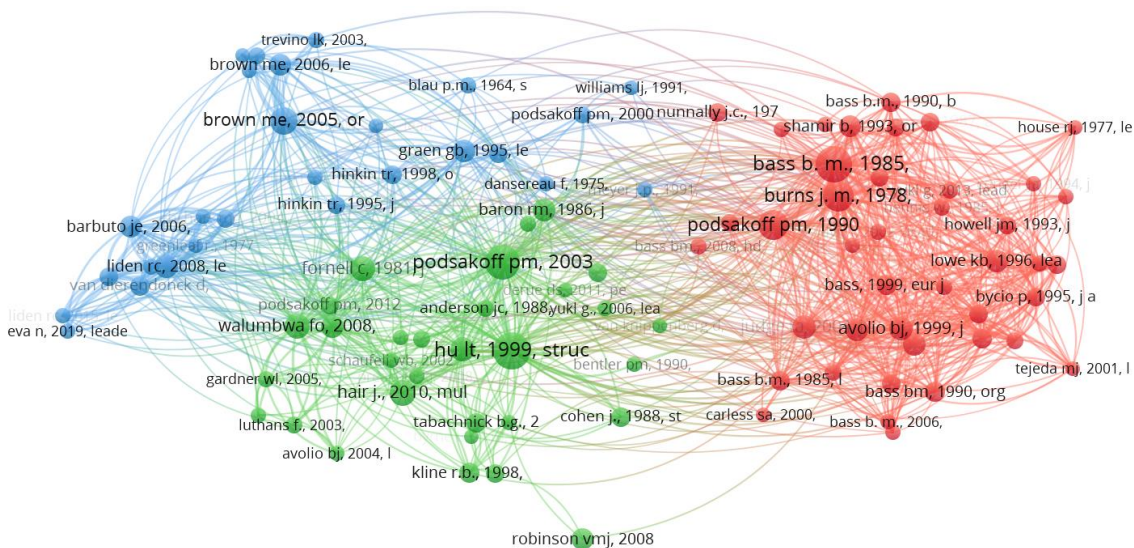
Visualization. To facilitate the cluster and strength analyses of the data downloaded from Web of Science, the researcher utilized VOSviewer software. Co-citation was chosen as the type of analysis, and cited references served as the unit of analysis. To ensure robustness, the threshold was set at a minimum of 20 citations.

The visualization yielded 117 main documents in four clusters. After careful examination of the documents and cluster themes, the researcher decided to streamline the list to the 100 most relevant documents, which ultimately resulted in three main clusters (see Figure 2.2 and Table 2.2). Table 2.2. presents the most significant documents, their main themes, and the corresponding link strength values.

Table 2.2. Cluster Information for Leadership Scale Documents

Clusters	Authors	Strength	Description of the Source
1) Theoretical Constructs	(Bass, 1985)	891	Transformational scale with five main factors: Charismatic leadership, contingent reward, individualized consideration, management by exception, intellectual stimulation
	(Burns, 1978)	694	Introducing two main constructs of transformational and transactional leadership
	(Podsakoff et al., 1990)	660	Measuring six transformational leader behavior: Articulating a Vision, Providing an Appropriate Model, Fostering the Acceptance of Group Goals, High Performance Expectations, Individualized Support, and Intellectual Stimulation How transformational leader behavior impacts follower's trust in the leader
	(Judge & Piccolo, 2004)	536	Testing the validity of transformation leadership in longitudinal research
	(Avolio, 1999a)	497	Re-examining the components of transformational and transactional leadership testing the six-factor model proposed by Bass (1985), using a larger and more heterogeneous sample
2) Validation methodology	(Hu & Bentler, 1999)	623	Cutoff criteria for fit indexes in covariance structure analysis: RMSEA, CFI, TLE, SMR, Gamma Hat, SRMR
	(Podsakoff et al., 2003)	679	Common method biases in behavioral research: a critical review of the literature and recommended remedies
	(Walumbwa et al., 2008)	453	Validation process of a theory-based measure for authentic leadership development
	(Avolio & Gardner, 2005)	393	Difference between authentic leadership, transformational leadership, servant leadership, spiritual leadership Components of authentic leadership: emotional contagion, positive social exchange, self-determination
3)?	(Brown et al., 2005)	452	Ethical leadership Trait and nomological validity Construct validity
	(Liden et al., 2008)	487	Servant leadership scale validation

Figure 2.2. Visual Map of Citation Networks and Clusters of 100 Documents (Leadership Scale)



Adaptive Leadership

Data Collection from Web of Science. A similar process was completed for adaptive leadership. A search was conducted in Web of Science using the keyword “adaptive leadership”, yielding a total of 1,518 documents. The researcher ran two rounds of refinement. In the first round, the results were filtered based on type of document (Article, review article, book chapter, paper proceedings) and relevant fields related to leadership (management, psychology, educational leadership, business, etc.). This refinement generated a list of 548 documents. Then, in the second round of refinement, the researcher went through the abstracts of these documents and removed documents irrelevant to leadership evaluation scale. This gave birth to the final list of 512 documents.

Visualization. The retrieved data from Web of Science needed to be visualized to do the cluster and strength analyses of documents. Using VOSviewer software to do network analysis, co-citation was chosen as the type of analysis and cited references as unit of analysis with

minimum number of 10 citations as the threshold. As a result, 51 main documents showed up in two clusters (see Figure 2.3. & Table 2.3.). Cluster 2 (green) was excluded from the review list as its contents primarily revolved around adaptive performance in the business context, which is not directly relevant to the focus of this research. Table 2.3 shows the most important documents with the highest link strength and their main themes.

Figure 2.3. Visual Map of Citation Clusters of 100 documents (Adaptive Leadership)

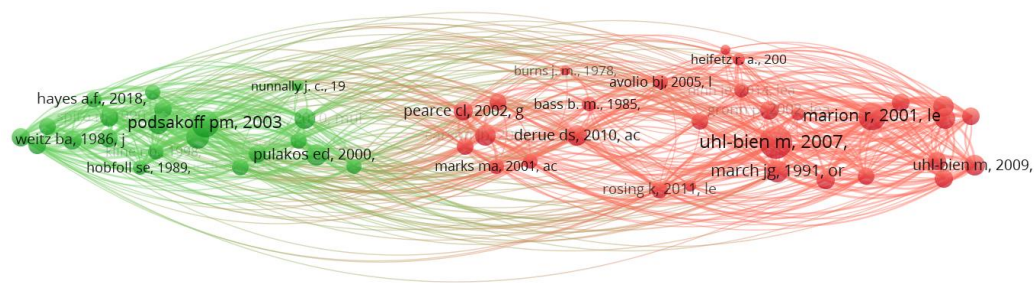


Table 2.3. Cluster Information for Adaptive Leadership

Clusters	Authors	Strength	Main Themes
1) Adaptive Leadership	(Uhl-Bien et al., 2007)	205	Discusses complexity leadership theory, a leadership paradigm that focuses on enabling the learning, creative, and adaptive capacity of complex adaptive systems within a knowledge-producing organizations
	(Marion & Uhl-Bien, 2001)	150	Develops a definition of organizational complexity Complexity and leadership science strategies for enabling complexity and effectiveness Relationship between complexity theory and important leadership theories.
	(Uhl-Bien & Marion, 2009)	88	How adaptive dynamics can work with bureaucratic functions to enable change Interaction between agentic behavior and non-agentic dynamics
	(Uhl-Bien & Arena, 2017)	69	Discusses three types of leadership: Operational, entrepreneurial, enabling leadership adaptive space conditions for leadership emergence Enabling leadership principles and practices Enabling leadership skills
	(Heifetz et al., 2009)	33	Practical guidelines for practicing adaptive leadership: Diagnosing the situation, managing self, engaging others, political landscape, intervening skillfully
	(Heifetz, 1994)	21	Definition of adaptive leadership Authority vs Leadership Biological foundations of adaptive leadership Adaptive vs technical challenges Adaptive leadership process

Discussion of the Literature

The bibliometric analyses unveiled prominent clusters of documents, including articles, books, and review papers. These clusters hold significant influence over the discourse on leadership development evaluation, scales, and adaptive leadership. These clusters served as gateways for reviewing the literature and contextualizing this research within the existing body of knowledge. In this discussion, the researcher aimed to highlight the patterns within these articles that inform the research, while also considering recent publications that might not have been captured in the bibliometric analysis.

Leadership Development Evaluation Clusters

The bibliometric analysis of the literature on leadership development evaluation revealed two primary clusters: paradigm shift and measurement approaches. The articles within the paradigm shift cluster focus on the evolution of human development theories and leadership theories. Specifically, they highlight the transition from traditional skill-based approaches to CDT (Kegan, 1982; Torbert, 2004).

The measurement approaches cluster, on the other hand, explores the various shifts in leadership development practices, particularly in relation to relational frameworks such as transformational-transactional leadership, shared/distributed leadership, and systems of adaptation (complexity leadership). These shifts can be traced back to seminal works by Burns (1978) and Bass (1985) and have continued through the more recent contributions of Mary Uhl-Bien and others.

The epistemological shifts manifested in the field of leadership development. Day (2000) wrote an important article at the turn of the century that highlights the difference between leader development and leadership development. He defined leader development as individualistic and

intrapersonal, while leadership as a social process and interpersonal. Leader development theories focus on human capital and individual skills like transformational leadership approach. In contrast, drawing upon collective leadership, leadership development suggests using relational systems and social capital to help build commitments in communities and organizations. Day (2000) argued that transformational leadership does not represent the thinking complexity needed to design and evaluate leadership development efforts. Introducing different practices of leadership development, Day (2000) emphasized the importance of research-based evaluation and assessment which he deems is usually overlooked.

Day (2000) informs this research in two main ways. First, his emphasis on linking intrapersonal to interpersonal skills in leadership development is a main concern of this research. Second, referring to Kegan (1994), he stressed complexity thinking needed in evaluating leadership development. While he did not elaborate on that in his 2000 article, he gave a clue to track this concept in his later works, especially Day et al. (2014) and Day and Dragoni (2015).

Day et al. (2014) is another important article that stood out in the bibliometric analysis. In this article, Day and his colleagues reviewed 25 years of advances in leader and leadership development and introduced different developmental theories and assessment methods. They gave importance to research and evaluation by asserting that “leadership development goes far beyond merely choosing a particular leadership theory and training people in behaviors related to that theory” (p. 63-64). Day et al. (2014) presented a long list of leader development theories (transformational, transactional, stage development theories) and longitudinal research conducted over the last three decades to measure variety of indicators as early predictors (e.g., intelligence, skills, experience, intrinsic motivation, adolescent, self-esteem, subclinical traits, positive

parenting) and outcomes like (leader performance, job performance, leadership role occupancy, leadership potential, team leadership, authentic leadership outcomes).

Day et al. (2014) further explored Kegan's CDT as a valuable approach for studying human development, particularly in the context of leadership. This perspective gained even more attention in the subsequent work by Day and Dragoni (2015). Distinguishing it from traditional skill-based approach and dynamic skill theory, Day and Dragoni (2015) referred to Kegan's orders of mind or consciousness as stages of development which offers a more holistic way to understanding self and environment. They argued that this approach to leader/ship development is "different from the forms of self-views (i.e., self-efficacy, self-awareness, and leader identity) offered as proximal outcomes of leader development" (p. 243). This is an important distinction that separates this research from KLC's 2021 evaluation study (Shams et al., 2023) using transformational leadership and self-views as the measurement tools.

Day and Dragoni's (2015) distinction between CDT and self-views made the researcher dive deeper into CDT literature. Following a progressive exploration, starting with Day and Dragoni (2015), the researcher delved into McCauley et al.'s (2006) review article. This exploration further led to the examination of the contributions of Kegan (1994) and Kegan and Lahey (2009). McCauley et al. (2006) did a literature review on how Kegan's (1982), Torbert's (1987), and Kohlberg's (1969) CDT frameworks were used in leadership. Among these three, their focus on Kegan's (1994) framework stood out for this research as they tried to see how three main orders of development/mind (dependent, independent, and inter-independent) were linked to leadership theories. McCauley et al. (2006) referenced Kegan's (1994) argument, which suggests that while transformational and transactional leadership theories call for independent and dependent leaders respectively, Heifetz's (1994) concept of adaptive leadership

necessitates the development of inter-independent leaders. According to McCauley et al. (2006), most of the research studying the relationship between Kegan's framework and leader effectiveness has focused on the first two rather than the latter (adaptive leadership).

Kegan and Lahey (2009) elaborated on the connection between their CDT framework and adaptive leadership. Drawing upon Heifetz's adaptive and technical challenges, they argued that solving complex challenges requires complexity of mind. In other words, to solve adaptive challenges individuals should move from socialized mind (dependent) to higher levels of mental complexity, self-authoring (independent) and transforming minds (inter-independent).

Kegan and Lahey's (2009) work informed this research theoretically as it explores the relationship between complexity of cognitive reasoning (through MRM) and ALB. Their insights provided a framework for understanding how cognitive complexity influences an individual's capacity to navigate and address adaptive challenges effectively. Moreover, Avolio et al.'s (2009) categorization of outcomes (Affective, Cognitive, Behavioral) concedes that this research can be situated within cognitive and behavioral evaluation approach. It also underlined the affective part mainly through the qualitative phase of the research.

Avolio (1999b) was another insightful piece that stood out for this research. Like Day (2000), at the turn of century, Avolio (1999) highlighted the importance of moving from individualistic leadership to more relational team and shared leadership. Referencing Heifetz (1994), Avolio (1999) discussed adaptive conflict and tension as the basis for building a highly developed team leadership system.

Scale Development Clusters

The bibliometric analysis of literature on leadership scale development resulted in two main clusters: Theoretical constructs and validation methods. In line with the paradigm shifts

discussed in the previous section, this literature review identified articles that were concerned about alignment of evaluation with the content of leadership development interventions both theoretically (Cluster 1) and methodologically (Cluster 2). Among these articles Avolio and Gardner (2005) and Walumbwa et al. (2008) came informative to this research. Besides similarities, Avolio and Gardner (2005) provided a clear distinction between authentic leadership and other theories and how this can be manifested in measurement tools. For instance, the authors show that while both authentic leaders and transformational leaders benefit from positive psychological resources, they are different in terms of ways they influence followers. This distinction provides insights into differentiating adaptive leadership from these two theories from a psychological standpoint. Notably, it is suggested that adaptive leadership relies more on negative emotionality (such as anxiety, heat, frustration, tension) (Heifetz et al., 2009; Shams et al., 2023; Uhl-Bien et al., 2007) rather than solely relying on positive psychological resources (such as hope, optimism, and self-efficacy) (Pellerin & Raufaste, 2020). This understanding can be particularly valuable in developing scale items to measure ALB in this research.

Adaptive Leadership Cluster

From Heifetz to Uhl-Bien

The bibliometric analysis of literature on adaptive leadership resulted in one main cluster. The concept of adaptive leadership, firstly, appeared in the literature as a teaching framework in the works of Ronald Heifetz in the late 20th and early 21st century as one of the manifestations of paradigm shift to collective leadership (Heifetz, 1994; Heifetz & Linsky, 2001; Heifetz et al., 2009). Later, adaptive leadership gained more attention in Uhl-Bien and her colleague's conceptualization of complexity leadership and the theory of complex adaptive systems (Uhl-Bien & Marion, 2001, 2009; Uhl-Bien et al., 2007; Uhl-Bien & Arena, 2017). Heifetz and Uhl-

Bien offer distinct perspectives on adaptive leadership, primarily concerning the locus and nature of leadership work. Heifetz focuses on individual as the initiator and mobilizer of relationships, whereas Uhl-Bien emphasizes the mutual processes within a system (such as a team or organization) that enable the emergence of leadership and navigate complexity. In sum, although both conceptualizations fall within the realm of collective leadership, they diverge in terms of the level of importance assigned to relational dynamics (Ospina, 2016). This differentiation is highly informative to this research, particularly in guiding the evaluation method and focus on the unit of analysis that is more relational competencies of individuals rather than mutual processes of a system.

Other Spots of the Literature

In this section, some of the recent relevant resources in the field of adaptive leadership, program evaluation, and scale development are highlighted and reviewed. This is essential because some of the important recent works could be overlooked in the refinement process due to not meeting the minimum citation threshold in the co-citation analysis.

Northouse's Adaptive Leadership Questionnaire

In his 2016 book, *Leadership: Theory & Practice*, Northouse suggested 30 items to assess individuals' ALB categorized in six dimensions: get on the balcony, identify the adaptive challenge, regulate distress, maintain disciplined attention, give the work back to people, and protect leadership voices from below. Using five-point Likert scale (Strongly disagree to strongly agree), Northouse suggested this questionnaire can be employed for practical applications using 360-degree and multirater evaluation approaches. According to Northouse (2016), this questionnaire was not designed for research purposes and its validity and reliability have not been established.

Raei's Adaptive Leadership with Authority Scale

Raei's (2018) dissertation is the most recent effort to provide an empirical foundation for adaptive leadership. Raei's work is distinct in several ways. First, he provided a comprehensive review and content analysis on adaptive leadership literature reflecting on the works of Heifetz and his colleagues (Linsky, Garshow, and Laurie), Mary Uhl-Bien and colleagues, Yukl and Mahsud (2010), Williams (2005, 2015), Northouse (2016), etc. He presented the similarities and differences between various conceptualizations of adaptive leadership and ultimately defined his own definition as "leadership that helps groups, departments, and organizations face difficult problems—problems that require a change in habits, values, or assumptions" (p. 79). Raei's work highlighted the confusion in the literature regarding the nature of adaptive leadership. He claimed, "The terms adaptive challenge, adaptive change, and adaptive leadership have been thrown around without paying attention to what Heifetz and his colleagues wrote" (p. 43). Comparing different scholarly works, he showed how adaptive leadership is being defined arbitrarily by different scholars as they pick one concept/dimension (adaptive challenge, learning needed, etc.) of it that fits their work.

Second, Raei's (2018) dissertation illustrated the difficulty of developing a scale rooted in adaptive leadership. Raei (2018) developed a leadership behavior scale with a focus on the authority aspects of adaptive leadership using seven dimensions: Distinguishing between adaptive and technical challenges, identify the stakeholders and their losses, create holding environment, regulate the distress, give the work back, protect voices of leadership without authority, use of self. In other words, Raei's scale measures the adaptive leadership behavior of those in authority. His initial multidimensional model was not confirmed through the validation process, and it led him to transform it into a unidimensional model with a good fit.

Adaptive Leadership & Student Development Programs

Weng and Hubbard (2022) framed adaptive leadership as a theory to evaluate the impact of case-in-point (CIP) courses in graduate leadership programs in an interview study. The authors argued that CIP is impactful on people's professional practice and levels of awareness. Also, using content analysis, Curan et al. (2013) assessed students' self and other learning by analyzing their reflection papers at the end of an adaptive leadership capstone course. The authors show that the adaptive leadership course was effective on students' openness to diverse perspectives, self-awareness, and relationship building. While the first study suggests incorporating CIP in leadership development programs, the second research recommends future research to study cognitive development as a result of adaptive leadership courses.

Adaptive Leadership & Adaptation-Innovation Theory

Highlighting limited empirical support for adaptive leadership, Seibel et al. (2023) made a theoretical connection between Heifetz's (1994) adaptive leadership and Kirton's (1976) adaptation-innovation (A-I) theory. They argued that A-I cognitive styles inform adaptive leadership holding environment. They suggest possible ways that A-I may contribute to adaptive leadership. They suggest that knowing people's cognitive styles (adaptive or innovative) may help managing disequilibrium in the groups and learning how people perceive challenges (adaptive vs technical). Regardless of whether Heifetz's conceptualization of adaptation aligns with Kirton's, this paper also highlights the importance of cognitive processes in facing adaptive challenges.

KLC's Your Leadership Edge & When Everyone Leads Books

Inspired by Chrislip and O'Malley (2013), KLC takes Heifetz's adaptive leadership concepts to civic arena by making the case for it in O'Malley and McBride's (2023) *When*

everyone leads and providing practical tools and strategies in O'Malley and Cebula's (2015) *Your leadership edge*. KLC argues that the gap between current reality and aspirations — discussed by Heifetz and Linsky (2002) — can be bridged by four main leadership competencies: diagnosing situation, managing self, energizing others, and intervening skillfully while adhering to five principles: 1) leadership is activity, not a position; 2) Anyone can do it, anytime and anywhere; 3) It starts with you and must engage others; 4) Your purpose must be clear; 5) It is risky. This conceptualization is the closest to origins of adaptive leadership framework even though there are differences in terms of scale (going beyond organizational context) and feasibility perceptions. Particularly, the second principle which seems different from how Heifetz et al. (2009) found leadership, “We find it extremely useful to see leadership as a practice, an activity that some people do some of the time” (p. 24).

KLC's 2021 Evaluation Study

This mixed-method research was conducted by KLC's Third Floor Research in collaboration with the LeAD Labs of Claremont Graduate University. Focused on KLC's Your Leadership Edge program, the quantitative phase measured the program's impact on participants' transformational leadership behavior and self-views right after the program, 30 days after, and one year after. The qualitative phase explained how the impact manifested in participants' professional life (Shams et al., 2023, 2024).

How the Literature Review Informed this Research

The literature review provided invaluable insights about how this research can be situated within the existing literature. Besides, the examination of the latest review articles — both in theory and methodology — indicated that many of these insights continue to be reiterated in the

literature (e.g., Crawford & Kelder, 2019; Vogel et al., 2021). The bibliometric literature review informed this research in four main ways:

- 1) It adopted a pragmatic-constructivist approach to study leadership development. Among different human development theories this research can be located under Kegan's CDT and Drath's (2001) interpersonal influence and relational dialogue knowledge structures. As such, regarding the primary outcomes for evaluation, it focused on cognitive and behavioral outcomes.
- 2) By focusing on individual's tendencies in building dyadic relationships and managing conflict, this research aligns with Heifetz's conceptualization of adaptive leadership. Moreover, it addressed the necessity of adaptive conflict in building relational leadership emphasized in Avolio (1999) as well as the connections between leader development and leadership development discussed in Day (2000) and Day and Dragoni (2015). In other words, this research underscores necessary individual competencies in developing relational collective leadership (not transformational and transactional).
- 3) This research follows the literature's suggestion about conducting longitudinal research to explore the cognitive processes involved in leadership development.
- 4) This research takes a distinct path by centering its focus on adaptive leadership and the collective leadership paradigm, as well as examining the cognitive outcomes of leadership development through the lens of CDT. In doing so, it explores a new direction that deviates from the predominant discourse and its authors mentioned as a primary concern in Vogel et al. (2021).

Furthermore, the targeted literature review primarily focused on the intersection of adaptive leadership, program evaluation, and scale development revealed the scarcity of

empirical and research-oriented evaluations on adaptive leadership programs within the existing body of literature. Moreover, it shows how the concept of adaptive leadership is being defined in inconsistent ways which motivated the researcher to adhere to Heifetz's works as the foundational source.

Also, it is important to note that this dissertation is different from Raei's (2018) dissertation as the only scholarly effort on developing a scale to measure adaptive leadership in terms of theoretical and methodological approaches as well as scope:

- 1) Theoretical Difference: First, this dissertation is not focused on the authority aspects of adaptive leadership. Grounded in the pragmatic foundation of adaptive leadership, it saw leadership as an activity that can be engaged with by individuals who may not hold formal positions of authority.
- 2) Methodological difference: The validation process went beyond CV, criterion validity, EFA, CFA, and reliability and tests concurrent, convergent, predictive, and discriminant validity of the scale to achieve higher validity and reliability. Moreover, this dissertation went through the process of validating adaptive leadership scales through the lens of CDT, specifically focusing on the perspective of mental complexity. Unlike a traditional theoretical framework, adaptive leadership lacks a comprehensive structure, making it challenging to isolate its components as distinct constructs. This research assumed that these behavioral statements can be organized according to the cognitive complexity required to effectively engage with them.
- 3) Scope: This dissertation went beyond the development of a scale and delved into a deeper exploration of the cognitive processes underlying adaptation and leadership behavior. To achieve this goal, the study validated additional scales, building upon

the established conceptual framework. Furthermore, these scales, along with the adaptation equation, were utilized in a program evaluation study focused on assessing the impact of KLC's 2023 path. This multifaceted approach aimed to provide a richer understanding of the intricate dynamics involved in adaptive leadership.

Theoretical Framework

This research studied the elements of leadership development process by focusing on the KLC's LDPs (YLE and LFC) as well as cognitive and behavioral developmental outcomes (Avolio et al., 2009; Day & Dragoni, 2016; Vogel et al., 2021). According to Avolio et al. (2009), "Cognitive outcomes represent how individuals perceive information, process information and make decisions while behavioral outcomes represent the actions taken by individuals that are observable" (p. 767). Inspired by this, the researcher's curiosity revolved around three broad inquiries that informed the research hypotheses and questions and consequently the theoretical framework: What are adaptive leadership actions? What are cognitive processes underlying them? how KLC's path programs are impacting these processes?

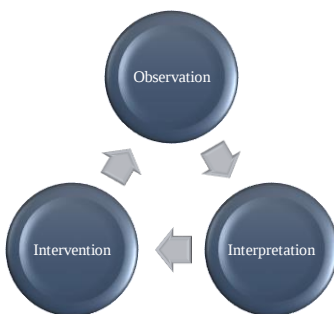
To address these inquiries, the research followed the constructive developmental tradition within human development theory. It particularly embraced Kegan's (1982, 1994) theory and its latest conceptualization encompassing cognitive complexity, adaptation, and three orders of mind: socialized mind, self-authoring mind, and self-transforming mind, as elucidated by Kegan and Lahey (2009). The recognition of Kegan's theory in the theoretical framework is primarily motivated by his emphasis on cognitive complexity and his argument regarding the need for a self-transforming mind in adaptive leadership.

Besides Kegan's theory as the overarching umbrella, this research benefited from three other theories to build its theoretical framework and provided an understanding of cognitive mechanisms underlying ALB: adaptive leadership framework in leadership, dual process theory (DPT) in cognitive psychology, and experiential learning theory (ELT) in adult learning. In this section, the researcher explained why these theories were selected and how they informed the theoretical framework to explore the research question and hypotheses.

Adaptive Leadership

In addressing the question of what constitutes adaptive leadership behavior, this research merely referred to Heifetz and his colleagues' definition as the point of origin and primary source due to the prevailing conceptual confusion that exists within the adaptive leadership literature (Northouse, 2016; Raei, 2018). Heifetz et al. (2009) defined adaptive leadership as “the practice of mobilizing people to tackle tough challenges and thrive” (p. 14). Inspired by evolutionary biology, adaptive leadership is about the change that enables the capacity to thrive, builds on the past, occurs through experimentation, and requires sustained periods of disequilibrium (Heifetz et al., 2009). Distinguishing leadership from authority, Heifetz and his colleagues have designed a practical way to make progress on tough challenges calling it adaptive work or response. Heifetz and Laurie (1997) suggested six principles for leading adaptive work: Getting on the balcony; identifying the adaptive challenge; regulating distress; maintaining disciplined attention; giving the work back to people; and protecting voices of leadership from below. This adaptive work is a cyclical process with three main steps: observation, interpretation, and intervention (Figure 2.4).

Figure 2.4. The Adaptive Leadership Process



Note. Adapted from (Heifetz et al., 2009, p. 54)

Step 1: Observation

Heifetz et al. (2009) highlighted the initial condition of a person for taking action and the natural intrinsic tendency to leap to action. They argued that people in organizations have a proclivity for quick intervention based on their intuition without assessing the situation. This is an internal automatic drive based on a default interpretation. According to Heifetz et al. (2009),

A default interpretation, leading to a default response ... can blind people to a wider array of solutions and ideas that might generate even more value. Also, a default that works in one setting, at one moment in time, may not necessarily work in another place or time. Because people tend to interpret new situations in ways that confirm the default, they fail to recognize the distinctive qualities of a new situation and thus cannot develop fresh solutions. (p. 95)

Heifetz et al. (2009) perceived adaptive challenges as difficult challenges,

... because their solutions require people to change their ways. Unlike known or routine problem solving for which past ways of thinking, relating, and operating are sufficient for achieving good outcomes, adaptive work demands three very tough, human tasks: figuring out what to conserve from past practices, figuring out what to discard from past practices, and inventing new ways that build from the best of the past. (p. 99)

Utilizing the analogy of stepping off the dance floor and onto the balcony for better observation, Heifetz and Linsky (2002) encouraged people to break free from their habitual patterns and gain a fresh perspective on the situation at hand. By getting distant from the action space, people can observe how their organizational, community, or social systems are responding to problems. This way you will get a new perspective on the system's structures, culture, defaults, power struggles, and reasons for work avoidance. Then, it is time to override these

defaults that may be felt as uncomfortable and risky (Heifetz & Laurie, 1997; Heifetz et al., 2009). Getting on the balcony is a pre-requirement for diagnosing the adaptive challenge.

Step 2: Interpretation

Separating adaptive challenges from technical ones is where Heifetz et al. (2009) said leadership begins. According to Heifetz (1994), adaptive challenges are different from technical ones with known solutions in the expert's or authority's hands. Unlike technical challenges, adaptive challenges can be addressed by going beyond authoritative expertise, case-specific learning, taking risks, and accepting change in priorities, beliefs, and habits. According to Heifetz et al. (2009), there are two main signals that show an organization is facing an adaptive challenge: cycle of failure because of applying technical solutions repetitively and dependence on authority. These two signals can be manifested in four archetypes:

- When there is a gap between organization's espoused values and behaviors
- When there are competing commitments
- When the unspeakable is not spoken
- When there is work avoidance (Heifetz et al, 2009).

In Heifetz's adaptive leadership, diagnosing the challenge has several aspects. Heifetz and Linsky (2002) introduced four diagnostic tools: distinguishing technical from adaptive challenges, finding out where people are at, listening to the song beneath the words, and reading the behaviors of authority figures. As part of what they called mobilizing the system, Heifetz et al. (2009) offered a framework for making multiple interpretations of the adaptive challenge. They argued that before working on the challenge, there needs to be a shift in people's interpretations, moving from simple to tougher interpretations.

Moreover, elaborating on the political thinking discussed in the 2002 book, Heifetz et al. (2009) discussed diagnosing the political landscape which involves identifying different stakeholders (factions) and their values, loyalties, losses, influence, and power. Heifetz et al. (2009) stressed the importance of having diversified interpretations and reasoning. They contended that the “capacity to orchestrate multiple interpretations of an organizational challenge is more likely to produce innovative insights than relying solely on one person’s viewpoint” (p. 167). Although this might take more time, friction, and pain, they believe that it makes addressing the adaptive challenge more efficient.

Step 3: Interventions

Effective interventions are designed and implemented based on multiple tough interpretations discussed earlier. Activities of the first two steps as well as the actions on change initiatives are considered interventions (Heifetz & Linsky, 2002). Therefore, one should get on the balcony, diagnose the challenge, and analyze and mobilize the factions as well as plan for three other main activities:

- Regulating/orchestrating the heat/distress/conflict through creating holding environment, controlling temperature, and pacing the work. This helps with orchestrating the conflict, tension, heat, and distress needed for the adaptive work and keeping it in a productive zone (Heifetz & Linsky, 2002; Heifetz et al., 2009). According to Heifetz (1994), “Holding environment consists of any relationship in which one party has the power to hold the attention of another party and facilitate adaptive work” (p. 104-105). This environment provides cohesion, collective attention, trust, framing, clarification, and focus needed for doing the adaptive work. Heifetz and Linsky (2002) described four ways to create a holding environment: A protected physical space with a facilitator to work

through the conflict, shared language or common history of the organization, deep trust in the institution, and rules and processes that give minority a voice.

- Give the work back by taking the work off your shoulders and place it where it belongs (Heifetz & Linsky, 2002).
- Hold Steady: by taking the heat, ripening the issue, and focusing the attention (Heifetz & Linsky, 2002)

Up to now, the main types of leadership behavior have been described and categorized into three main steps of observation, interpretation, and intervention cycle. Heifetz and his colleagues talk about a gap between defaults and new ways (behaviors) that can be bridged by this adaptive process of disturbance, diagnosis, and intervention. Heifetz et al. (2009) declared that this process takes you out of your daily routine and, therefore, it is risky. Running this adaptive work requires “will plus skill”,

If leadership involves will and skill, then leadership requires the engagement of what goes on both above and below the neck. Courage requires all of your heart, mind, spirit, and guts. And skill requires learning new competencies, with your brain training your body to become proficient at new techniques of diagnosis and action. You might think about this idea as the convergence of multiple intelligences (intellectual, emotional, spiritual, and physical) or the collaboration among physical centers (mind, heart, and body). But the central notion is the same. Your whole self-constitutes a resource for exercising leadership. (p. 61)

In another part, Heifetz et al. (2009) discussed orienting purposes as drivers of adaptive work,

There is no reason to shoulder the difficult work of leadership if you do not have compelling, higher purposes to serve, whether saving the world, renewing your

organization, or helping your community meet long-standing challenges and thrive through tough times. Your purposes provide the inspiration and the direction for your actions. (p. 304)

What these quotes suggest is that skills (new leadership competencies discussed in the cycle) and will (higher purposes) are two essential elements of adaptive leadership performance. Now, the following question for building the theoretical framework is *what is the cognitive mechanism behind it?* What are the cognitive processes involved in disturbing the observation, making multiple interpretations, and intervention? And what are the cognitive processes involved in activating the higher purposes? In response to these *immediate cause questions* (Tinbergen, 1963), DPT was included as the second theory in this theoretical framework.

Dual Process Theory of Reasoning (DPT)

To measure cognitive complexity, researchers mostly have solicited Kegan's subject-object qualitative interviews. However, the relationship between cognitive complexity/orders of mind and leadership behavior has rarely been touched. DPT gives a framework to study and measure the orders of mind required in adaptation quantitatively. If we conceive leadership behavior as a rational action (as opposed to a-rational not irrational), DPT could describe the reasoning processes involved in leadership behavior. While adaptive leadership framework tells us performance is about skills of disturbance, interpretation, and intervention directed by higher purposes, DPT explains the mechanisms underlying it.

DPT assumes that humans use two different processing systems or types of minds. One is automatic, peripheral, impulsive, and fast that relies on well-learned information and has roots in human's routine activities, habits, and recurring tasks of daily life. The other is slower, more deliberative, analytic, cautious, lazy, and slow that relies on rules and symbolic logic (Evans,

2008; Schneider & Shiffrin, 1977; Stanovich, 1999, 2008; Kahneman, 2011; De Neys, 2018).

This theory has been used in several fields such as communication, political science, and leadership explaining the cognitive processes involved in persuasion (Elaboration Likelihood Model) (Petty & Cacioppo, 1986), decision-making (Marcus, 2002), and leader-follower influence (Kelman, 1958). The following section presents the latest extension of DPT and describes how its processes function.

Tripartite Model of Mind

DPT received a new extension by Stanovich (2008, 2011). First, in response to criticisms (e.g., vague definitions, misalignment of the attributes, unconvincing evidence about the existence of two systems; Evans & Stanovich, (2013)), it was clarified that the features of the two types of processes are descriptive rather than defining. Introducing *working memory* and *cognitive decoupling* as two main defining features that separate *Type1* from *Type 2*, Stanovich (2008) dove deeper into Type 2 and proposed two levels of mind within it: *algorithmic* and *reflective* minds. According to Stanovich (2008), the algorithmic mind is a level of processing that “implements the long-leashed goals of the analytic system. The algorithmic system is subordinate to epistemic thinking dispositions that exist at the reflective level of processing [which] regulates behavior at high level of generality” (p. 57). These reflective processes “operate to set the goal, agenda or they operate in the service of epistemic regulation. Such processes that set and regulate the goal and epistemic agendas” (Stanovich, 2011, pp. 66-67) need to be engaged in executive function tasks (rational action and behavior).

According to the recent version of DPT, in many of our daily routine activities, we use our Type 1(The Autonomous Set of Systems-TASS) to respond and act like expressing an over-practiced opinion. When the automatic system’s response to a new situation is not available or

enough, Type 2 as a default interventionist (Evans & Stanovich, 2013) or parallel process (Kahneman, 2011) starts operating. This initiation requires surprise, anxiety, or dissonance to happen. At this moment, the reflective mind sends a call to the algorithmic mind for decoupling a secondary representation from a primary one that gives the person the ability to manipulate it and create a new alternative. Overriding Type 1 response is called cognitive decoupling 1 (or inhibition as the executive function) and simulating an alternative is called cognitive decoupling 2 (or updating as the executive function). If this operation which Stanovich (2011) called *hypothetical thinking* continues and repeats by the algorithmic mind, it is called cognitive decoupling 3 (or shifting as the executive function). In sum, the reflective mind is the initiator of cognitive decoupling (1,2,3) and the algorithmic mind is the operator.

DPT in Leader Development

Day et al. (2009) reflected on the importance of these levels of mind in building an integrative leadership development theory and suggests pairing cognitive processes with skills and abilities needed for effective leadership. Referring to Fischer et al.'s (2008) dual process model and Kitchener's (1983) tripartite model of thinking, Day et al. (2009) highlighted the importance of critical thinking, reflective judgement, and epistemic cognition in leader development. They discussed that these cognitive processes are needed in tackling ill-structured problems that do not have best solutions and cannot be resolved with certainty. According to Day et al. (2009),

Cognitive demands in complex novel and ill-structured problems would seem to require mature epistemic cognition, reflective judgement, and well-developed critical thinking skills. Leaders cannot rely solely on prior training and experience to solve such problems. These aspects of cognitive development should be a significant focus as part of long-term

leader development across all of the primary venues, including formal education, on-the-job assignments, and self-directed development. (p. 102)

Day et al. (2009) claimed that epistemic cognition, along with leader identity, is a potential lever for leader development. They suggested a framework to theorize cognitive processes in leader development that puts epistemological beliefs (e.g., epistemic cognition, reflective judgement) at the highest level of thinking that influence cognitive functions like decision making and information processing (critical thinking) at the basic level of thinking. While Day et al.'s (2009) *ill-structured problems* communicate Heifetz's adaptive challenges, the cognitive processes described above resonate with Stanovich's (2011) tripartite model that conveys epistemic dispositions at higher order of thinking command algorithmic processes of mind. While Day et al.'s (2009) epistemic cognition may exhibit some variations in conceptualization (not in level of mind) compared to Stanovich's (2011) epistemic disposition, they both align in recognizing critical thinking as an analytic/algorithmic skill.

The connections between Day et al.'s (2009) and Stanovich's (2008, 2011) DPT becomes stronger when we notice self-regulatory cognitive processes are highlighted in Day et al.'s (2009) integrative leader development theory. Day et al. (2009) introduced five particular aspects of self-regulation as important individual differences in leader development: Self-regulatory strength, goal orientation, generalized self-efficacy, self-awareness, and implementation intentions. Day et al. (2009) suggested that "these variables influence or moderate the process of leader development" (p. 211). Similarly, Stanovich (2011) prepared a long list of fluid rationality and the reflective mind components as individual differences affecting rational action including: self-control skills, sense of self-efficacy, goal hierarchy, controlled emotional regulation.

This connection more strongly justifies that DPT can be replicated to the field of leader/ship development. However, can we connect DPT to a particular framework like adaptive leadership?

Connections Between Adaptive Leadership and DPT

In Stanovich's perspective, reflective and algorithmic minds are the two main factors of human rationality. It appears that these two factors communicate Heifetz et al.'s (2009) skill and will as two requirements of adaptive work. Particularly noteworthy is Stanovich's (2011) assertion that reflective processes are equivalent to what is commonly referred to as the "will" in folk psychology (p.82). Furthermore, the operator role of algorithmic and the commander role of reflective mind (Stanovich, 2011) resonates with Heifetz et al. (2009) saying that "Your purposes provide the inspiration and the direction for your actions (p. 304)."

The connections between adaptive leadership and DPT can also be seen in their operational processes. One can compare the constructs of dissonance, anxiety, and inhibition (Cognitive Decoupling 1) in DPT with disturbance of defaults in adaptive leadership; making multiple interpretations and interventions in adaptive leadership with updating and shifting simulations as executive functions (Cognitive Decoupling 2 & 3). In other words, it appears that the three types of cognitive decoupling are the cognitive mechanisms for the skills of disturbance (balcony observation), interpretations, and interventions. These cognitive mechanisms are enabled by the epistemic dispositions (reflective mind) that underlie an individual's will.

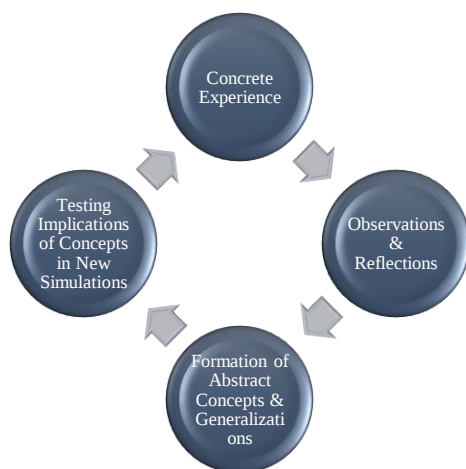
However, the researcher could not connect these two theories/frameworks merely based on conceptual similarities. To demonstrate the validity of this relationship, a connector theory is needed: Experiential Learning Theory (ELT). This theory provides a deeper explanation of the

connection between DPT and adaptive leadership by responding to this ontogeny question (Tinbergen, 1963) that *how has the adaptive behavior evolved during human's lifetime?*

Experiential Learning Theory (ELT)

ELT is an adult learning and educational theory that stresses the importance of experience and reflection in the learning process. Kolb and Fry (1975) introduced experiential learning model as the most seminal of the contributions that group dynamic scholars like Kurt Lewin made. Based on Lewin's model, change and growth happen through a cyclical process with four steps "1) here-and-now experience 2) collection of data and observations about the experience 3) analysis of data and conclusions and giving feedback to the actors in experience 4) modification of actors' behaviors and choices" (Kolb & Fry, 1975, p. 34) (Figure 2.5). This basic model of growth formed the foundation of Kolb's ELT.

Figure 2.5. The Lewinian Experiential Learning Model



Note. Adapted from (Kolb & Fry, 1975, p. 33)

Four Adaptive Modes of Learning

Inspired by Lewin's model as well as Dewey's principles of experience (continuity and interaction) (1938) and Piaget's cognitive development stages, Kolb & Fry (1975) argued that an effective learner needs four abilities: Concrete Experience (CE), Reflective Observation (RO), Abstract Conceptualization (AC), and Active Experimentation (AE). Kolb (1984) said "new knowledge, skills, or attitudes are achieved through confrontation among [these] four modes of experiential learning" (p. 30):

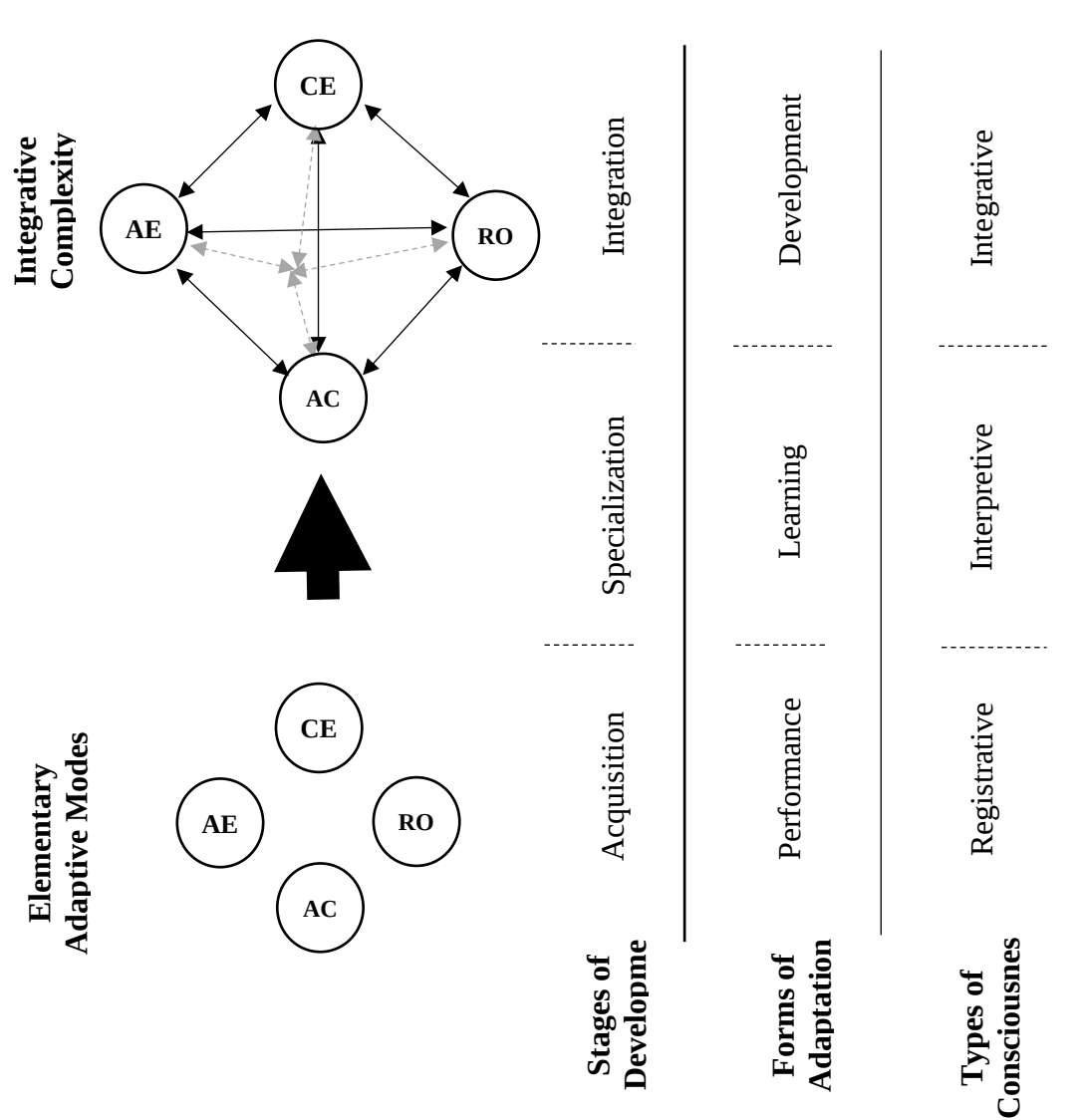
- Apprehension: Concrete experience (CE) of our immediate environment
- Comprehension: Abstract conceptualization (AC) and grasping the experience (Building a general theoretical model)
- Intention: Reflective observation (RO), Internal reflection-Internal manipulation (building multiple models)
- Extension: Active experimentation (AE), active external manipulation

Kolb (1984) described learning as a conflict-filled process. Kolb (1984) argued, "the structural bases of the learning process lie in the transactions among these four adaptive modes [and] knowledge results from the combination of grasping experience and transforming it" (p. 41). According to Kolb (1984), people may have special abilities in each of these learning modes. However, "more powerful and adaptive forms of learning emerge when these strategies are used in combination" (p. 65). Yet, this complete emergence and integrative complexity is not without difficulty. According to Kolb (1975, 1984), prehension modes (CE & AC) and transformation modes (RO & AE) are usually in battle. While CE and AC may inhibit each other, RO & AE may stop each other. This conflict makes it hard to reach to higher levels of learning.

Moving from Kolb's (1984) elementary modes of learning toward integrative complexity happens in three stages (Figure 2.6) (Kolb, 1984). First, Acquisition (self as undifferentiated-immersed in the world); Second, specialization (self as content-interacting with the world); Third, integration (self as process-transacting with the world). Referring to Piaget's cognitive development theory, Kolb and Fry (1975) believed that up to age of 15, humans gain the basic cognitive structures of all four abilities (Acquisition). After the age of 15, the development of these four abilities (moving to specialization and integration) depends on the choices (living, learning, socialization, culture, etc.) a person makes. Moving between these stages requires three levels of adaptation: *performance, learning, and development*.

Learning adaptations are respectively governed by registrative, interpretive, and integrative consciousness. The three types of consciousness occur differently in an individual's lifetime. Registrative consciousness happens when the existing model of a concept is affirmed or extended by the experience. This happens through the "performance" form of adaptation. If the experience changes the existing model, it means the consciousness took an interpretive form that alters, improves, and creates alternatives for the existing model. This second consciousness governs the "learning" form of adaptation. Kolb (1984) said, at this level "one develops the ability to observe experience from different perspectives" (p. 156). The third level of consciousness (The most complex one) sets the purpose and focus for choosing a new experience in a particular direction based on the choices given by interpretive consciousness. According to Kolb (1984), "This level begins with a development of an experimental, hypothesis testing approach to action that introduces new tentativeness" (p.150). While interpretive consciousness is more analytic, integrative one is more synthetic.

Figure 2.6. From Elementary Modes to Integrative Complexity in Experiential Learning



Note. Adapted from Kolb (1984).

ELT as the Connector of Adaptive Leadership and DPT

Connection to Adaptive leadership

The adaptation process in ELT, whether utilizing Lewin's model (observation, analysis, modification) or Kolb's model (CE, RO, AC, AE), is harmonized with Heifetz's adaptive leadership cycle (observation, interpretation, and intervention). Kolb (1984) defined learning as a holistic process of adaptation to social and physical environments. He stated that "... the failure to modify ideas and habits as a result of experience is maladaptive" (p. 26). He named the aforementioned cyclic modes of learning as adaptive processes. The connection between ELT and adaptive leadership becomes clearer considering Kolb's (1984) assertion that ELT extends beyond just learning and encompasses more adaptive concepts such as problem-solving and attitude change.

Another way to connect ELT to adaptive leadership is through Kegan's CDT. Although Kolb's (1984) development stages (Acquisition, specialization, integration) has more emphasis on learning processes, his framing of cognitive and consciousness levels communicate Kegan's stages of development, orders of mind, and cognitive complexity needed for adaptive leadership: socialized mind, self-authoring mind, and self-transformative mind (Kegan, 1982; Kegan, 1994; Kegan & Lahey, 2009). The connection between these two conceptualizations of human development will be discussed later in this chapter.

Connection to DPT

While ELT's one hand is connected to adaptive leadership, its other hand is linked to DPT. ELT establishes its role as a connector theory firmly when Kolb (2014) made direct reference to DPT. Kolb replaced Roger Sperry's two-hemisphere theory as the supporter of his ELT in 1984 with a new cognitive theory. Interestingly, Kolb (2014) made specific reference to

Kahneman's (2011) and Evans' (2008) DPT. These two psychologists played a crucial role in shaping Stanovich's (2008) DPT and his Tripartite Model of mind (2008, 2011). Kolb (2014) said,

Recent fMRI research has added substantially to our understanding of James' dual knowledge theory which formed the foundation for contemporary work on dual processing (Evans, 2008), made popular by Kahneman's recent book *Thinking, Fast and Slow* (2011) describing apprehension as system 1 and comprehension as system 2. (p. 87)

In fact, Kolb (2014) described the previously mentioned confrontation between four modes of learning (apprehension with comprehension and intention with extension) as the conflict between Type 1 and Type 2 of the brain. In other words, running an adaptive learning process requires Type 2 of the brain. Knowing the alignment of ELT's adaptation process and adaptive leadership cycle, once again adaptive leadership can be connected to DPT. Therefore, it can be concluded that ELT plays a connector role between these two theories. From the research perspective, this connection suggests that DPT can be used to model the cognitive processes involved in adaptive leadership behavior.

Furthermore, it seems cognitive decoupling 1, 2, and 3, run by the reflective mind, are involved in adaptation governing mechanisms (Table 2.4). Also, Kolb's (1984) types of consciousness as adaptation governors — especially interpretive and integrative — appears to communicate Stanovich's (2008, 2011) cognitive decoupling 2 and 3. Interpretive consciousness is about alerting the system about the changes needed which speaks to cognitive decoupling 2 which runs the process for creating alternative scenarios for an existing situation. Furthermore, integrative consciousness involves establishing a clear purpose for action, which aligns with the concept of decoupling 3. Decoupling 3 refers to the process of sustaining simulation based on the

purposes and dispositions that have been adopted. Although testing this hypothesis is not the aim of this research, this conceptual assumption helped the researcher with explaining the connections between theories.

Theoretical Framework Components

Concepts and Constructs

Figure 2.7 shows how DPT, ELT, and adaptive leadership speak to each other. These three frameworks imply a gap between status quo and new future using different terms. They describe a conflict-filled process to bridge this gap: DPT (battle between Type 1 vs Type 2), ELT (conflict between prehension vs transformation), and adaptive leadership (disequilibrium between stability and change). This overall communicates the tensions between socialized, self-authoring, and self-transforming minds in CDT.

The concept of adaptation, in response to these tensions, serves as the primary link between these theories. Heifetz and his colleagues claimed that there is a gap between defaults and new ways that can be bridged by an adaptive process of disturbance, diagnosis, and experimentation (skills) directed by higher purposes (will). DPT highlights this gap between routines/habits and new behaviors that can be spanned by three executive functions (Cognitive Decoupling 1, 2, 3) as the functions of adaptation (algorithmic mind + reflective mind). Also, ELT implies a gap between habits and modified learning that can be bridged through the adaptive modes of learning (Prehension + Transformation).

This theoretical framework provides a lens to study adaptation and leadership development at three levels: cognitive, learning, and action level. It can theoretically inform us what cognitive processes are involved in learning and practicing leadership and how teaching intervention at the cognitive level can impact learning and behavior. In other words, it can

inform us how LDPs' content can target cognitive processes (what) involved in rational action and experimentation.

Variables

This research studied the cognitive processes involved in exercising adaptive leadership (action level) and gave recommendations for the learning level. First, in the quantitative summative phase, it developed measurement tools to evaluate participants' ALB and the cognitive processes involved in the adaptation process (Reflective Mind). Informed by the theoretical framework, below are the variables that can measure cognitive processes underlying algorithmic and reflective levels of mind and the adaptive leadership behavior as rational action.

Adaptive Leadership Behavior (ALB)

As mentioned earlier, here leadership behavior is defined as a rational action (as opposed to a-rational, not irrational). The items of these scales were constructed based on adaptive leadership literature (Solely Heifetz's work).

Adaptive Leadership/Adaptation and Leadership Self-Efficacy (ALSE)

Defined in the Introduction chapter. According to Stanovich et al. (2011) self-efficacy is one of the components of fluid rationality that "may require sustained decoupling and thus implicate the computational capacity of the algorithmic mind" (p. 194). Day et al. (2009) also conceptualized self-efficacy as one of the aspects of self-regulation that "has been included in higher-order constructs addressing one's overall assessment of the self" (p. 196). Hannah et al. (2012) also highlighted the importance of separating action and self-regulation aspects of self-efficacy. In sum, while ALSE is considered mainly as a reflective mind component, it includes algorithmic mind processes as well.

Adaptive Leadership/Adaptation and Leadership Self-Awareness (ALSA)

Informed by Heifetz's adaptive leadership, a new self-awareness scale is developed based on Taylor's (2010) two components: 1) awareness of self-character, traits, assessment of strengths and weaknesses, sense of purpose or calling, core values, and motivations and desires 2) the leader's ability "to anticipate the views of others" in relationship with them; "becoming aware of influence through impulses to self-focus" (pp. 57-58).

Master Rationality Motive (MRM)

Stanovich's (2011) MRM is used in this study to measure the reflective mind. Inspired by Velleman (1992), Stanovich (2011) defined MRM as "the motive that drives the search for rational integration across our preference hierarchies" (p. 81). "As a thinking disposition that needs to be assessed under typical, not maximal, conditions, [MRM] measures variation in the desire to act in accordance with reasons" (Stanovich, 2011, p. 94). MRM indicates people's higher-order tendency to prefer their preference for exercising leadership and resist the preference of not exercising leadership. Stanovich (2008) argued that this tendency "to prefer a preference" is something that cannot be measured by intelligence tests like IQ. MRM measures the tendency to decouple particularly inhibition (decoupling 1) and updating (decoupling 2). It can be considered as a measurement of tendency to think in a more complex way (i.e., creating metarepresentations of the situation; Stanovich, 2011). The higher the MRM, the more skills an individual might possess in making metarepresentations.

Adaptive Leadership/Adaptation and Leadership Shifting-Reflectiveness (ALSR)

This scale captures the thinking dispositions that transform rational integration into action (Stanovich, 2011). While MRM measures tendencies toward forming and changing beliefs, ALSR measures goals and purposes of exercising leadership. The definitions of ALSR and

MRM speak to Day et al.'s (2009) goal orientation construct as one of the aspects of self-regulation in leader development.

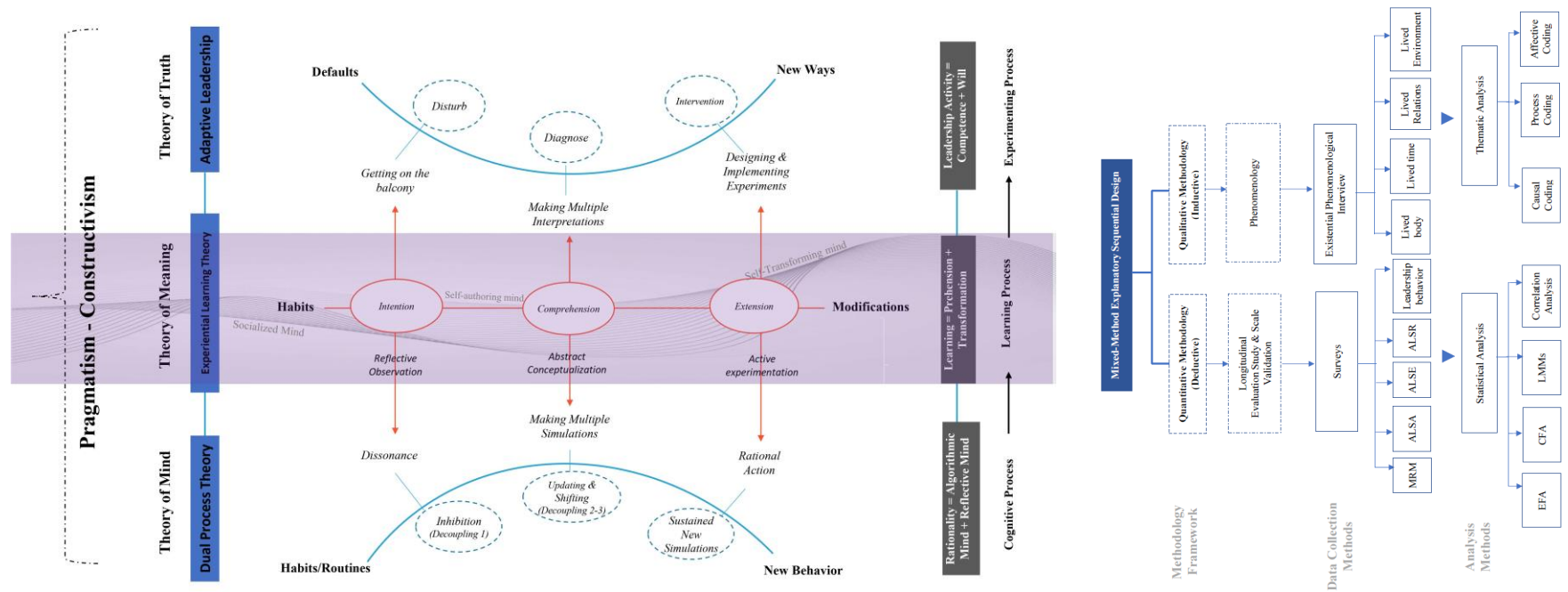
The relationship between these scales was theorized based on DPT and its Tripartite Model of Mind. The two primary components of this phase examined the occurring leadership actions and the associated cognitive processes, including algorithmic and reflective minds. The development of scales identified the pertinent and measurable thinking dispositions related to adaptive leadership. Furthermore, the evaluation study demonstrated the extent to which participants' different levels of mind (Stanovich, 2011) and orders of mind (Kegan & Lahey, 2009) are engaged in adaptive leadership experiments and whether KLC is improving them.

Identifying adaptive leadership behaviors and the cognitive processes underlying them can inform teaching content and methods. This can help KLC modify its course content and teaching methods in 2024. To make this happen, the researcher added a formative qualitative phase to the quantitative phase of the research. The qualitative phase (interview study) aimed to translate the quantitative findings into practical teaching recommendations by exploring how these findings manifest in real-world experiences. For example, if the quantitative phase indicated the presence of a reflective mind that drives individuals towards new actions, the qualitative phase would delve into the "how" of this mechanism within real situations. Uncovering this "how" could provide valuable insights to KLC's teaching team, offering potential strategies for addressing and cultivating the reflective mind.

Table 2.4. Theoretical Connections Between Adaptive Leadership, DPT, CDT, and ELT

Kolb's types of adaptation	Kolb's (1984) stages of development	Kolb's types of consciousness	Kegan's cognitive complexity level (orders of mind)	Stanovich's Cognitive Decoupling	Heifetz's Adaptive Leadership
Performance	Acquisition	Registrative	Socialized mind	Resisting TASS-Disrupting performance	Disturb
Learning	Specialization	Interpretive	Self-authoring mind	Creating new simulations/alternative of the existing model (Hypothetical thinking)	Diagnosing the situation (making multiple interpretations)
Development	Integration	Integrative	Self-transforming mind	Choosing a simulation and driving action based on thinking dispositions (e.g., purpose, epistemic values, goals)	Designing experiments (Intervention)

Figure 2.7. Theoretical/Conceptual Framework



Creating a Mapping Tool for Tracking Cognitive Complexity

To make sense of adaptation and see at what level of adaptation a person resides, an integrated framework of Kolb (1984), Kegan and Lahey (2009), and Stanovich's (2011) adaptation levels (Figure 2.8) was developed in this research. Using this mapping tool, it became possible to assess the extent to which KLC's LDPs contribute to enhancing individuals' cognitive complexity.

Level 1: Acquisition/Socialized Mind

This is the performance level of adaptation. The initial stage of cognitive development in which people act based on what they receive from outside. The mind is at its simplest level of complexity. In this stage, there is a direct relationship between CE and AE (Kolb, 1984). Kegan and Lahey (2009) stated "we are shaped by definitions and expectations of our environment" (p. 17). In this phase, people build a representational model for their mission and have a registrative consciousness (Kolb, 1984; Stanovich, 2008). They try to build an acceptable face by showing alignment with others especially important and more influential ones (Kegan & Lahey, 2009). Center for Creative Leadership (CCL) calls this stage as dependent-conformer (McGuire & Rhodes, 2009)

Level 2: Specialization/Self-Authoring Mind

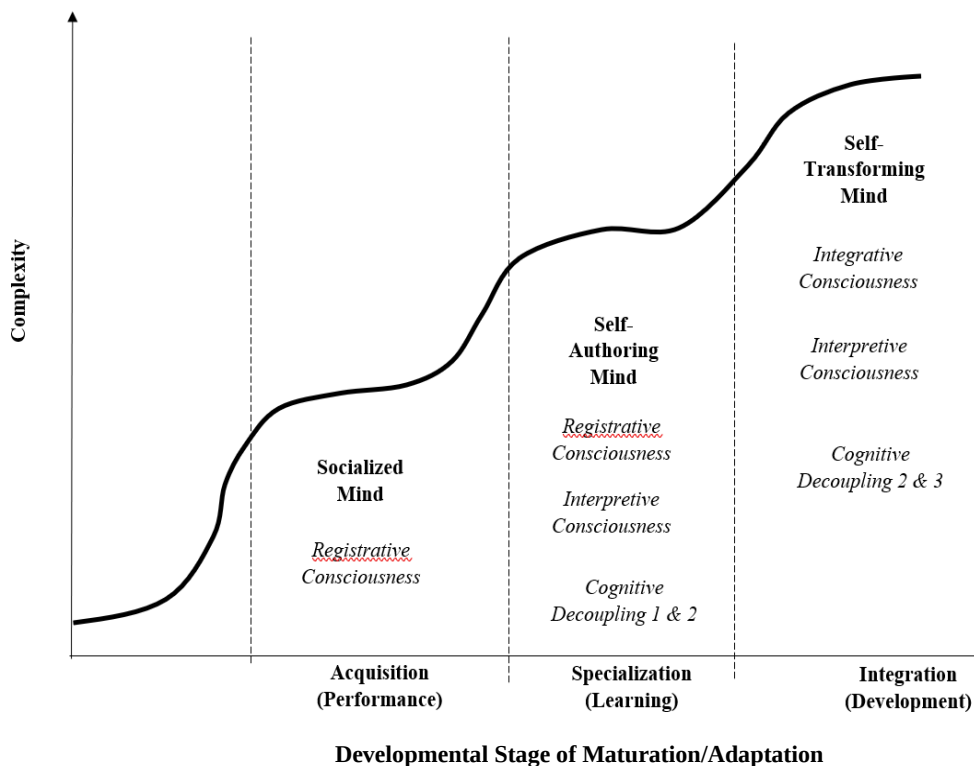
In this level, specialization happens as a result of learning. According to Kegan and Lahey (2009), in this phase a person "has a direction, agenda, stance, strategy, and analysis of what is needed... it has a filter for what it will allow to come through" (p. 19). In this stage, registrative consciousness is in place and interpretive consciousness might be initiated to create alternative models for the existing model (Kolb, 1984). From Stanovich's (2011) cognitive standpoint, this is the phase that cognitive decoupling 1 and 2 happen and a person tries to

evaluate their actions. However, there is only one recognized perspective. To make it multiple, one should proceed to the third stage. Center for Creative Leadership (CCL) calls this stage as independent-achiever (McGuire & Rhodes, 2009).

Level 3: Integration/Self-Transforming Mind

In this stage, development happens (Kolb, 1984). Kegan and Lahey (2009) asserted that, in this phase, people stand back and make space for modifying their agenda. “They seek information that may lead them or their team to enhance, refine, or alter the original design to make it more inclusive” (p. 20). Both Kolb (1984) and Kegan and Lahey (2009) claimed that learning and development in this phase happen through a transactional relationship with the environment. Stanovich’s (2011) cognitive decoupling 2 and 3 and Kolb’s (1984) interpretive consciousness occur in this phase meaning that multiple simulations and alternatives for the existing model are created. Interpretive consciousness will be followed by integrative consciousness which integrates alternatives based on personal/organizational beliefs, purposes, and priorities (Kolb, 1984). Stanovich (2011) called this rational integration. Unlike the second phase which puts a filter on what it receives as feedback from others, those with self-transforming minds receive whatever comes to their doors (Kegan & Lahey, 2009). Center for Creative Leadership (CCL) names this stage interdependent-collaborator (McGuire & Rhodes, 2009).

Figure 2.8. Adaptation and Complexity Stages



Note. Adapted from Kegan & Lahey (2009), Kolb (1984), and Stanovich (2011)

Ontological/Epistemological Commitments

This mixed-method research and its conceptual framework is situated at the intersection of pragmatism and constructivism. It is constructivist “in the sense that it deals with a person’s construals, constructions, and interpretations” (McCauley et al., 2006). It is pragmatic since these interpretations stem from the meaning a person makes of a leadership experience and action (Pierce, 1878).

This research followed pragmatic tradition particularly in the quantitative phase. In this phase, there are three major hypotheses that tested in the KLC’s 2023 path evaluation study. This hypothesis testing did not happen through objective fact-finding but by focusing on the individual’s self-report of their leadership experiments. It did not take a verification process, but

a falsification and abduction process. In this view, there is a reflective relationship between theory and practice. Therefore, the theory takes a Darwinistic evolutionary process (Van de Ven, 2007).

In the sequential qualitative phase, however, the researcher conducted one-on-one existential interviews without any theory in mind and focused on the participants' stories of leadership experiments. He particularly focused on concrete experiments without asking for any analytic reflection. Doing so, he searched for evidence that either supported the results of the first phase or explained the surprises. Both of these pieces of evidence took him deeper into the reasoning processes involved in leadership experiments and provided insights for the KLC's teaching team.

In sum, the researcher employed three theories of mind, meaning, and truth based on Meyers' (1999) pragmatism. DPT is the theory of mind that posited the hypotheses and ideas of rational action. ELT is the theory of meaning that revealed the connection between theory of mind and action. Finally, adaptive leadership is the theory of truth which showed whether the hypotheses are true in guiding action and prediction. Any surprise arising from the theory of action necessitates modification in the theory of mind.

Highlights of the Chapter

- Systematic literature review was conducted in three areas: Leadership development evaluation, leadership evaluation scale, and adaptive leadership.
- Theoretical framework was developed weaving three main theories through the common concept of adaptation: adaptive leadership, ELT, and DPT
- A theoretical framework diagram was created and illustrated encompassing concepts, constructs, and variables.

Chapter 3 - Methodology

Evaluating LDPs has been one of the popular research areas in leadership development. However, addressing the distance between learning and practicing leadership from the cognitive reasoning standpoint has received relatively limited attention (e.g., Avolio & Hannah, 2008; Day et al., 2009; Hannah, 2006). Working toward this goal, this research is anchored in this idea that LDPs should be tailored to people's cognitive reasoning process. Being aware of cognitive processes of learning and acting can inform LDPs to be more targeted and purposeful in designing modules. To demonstrate this, the researcher used KLC's 2023 path as the strategic research site/material to conduct a mixed-method study. This research has three main objectives that serve toward the main aim of the research "To develop valid, reliable, generalizable, and practice-reflective assessment model for adaptive leadership":

- 1) Developing Instruments for Evaluating LDPs with Adaptive Leadership Approach
- 2) Measuring the Effectiveness of KLC's Leadership Development Path 2023
- 3) Providing KLC's Teaching Modules with Specific Recommendations to Improve the Effectiveness of the LDPs

Hypotheses/Questions

The objectives, hypotheses, and questions of the research revolved around the three main challenges of LDP evaluation discussed above:

Objective 1: Developing Instruments for Evaluating LDPs with Adaptive Approach

This objective was defined to address the *teaching-assessment* and *measurement gaps* and has one main hypothesis comprising of four subsections.

Hypothesis 1: Adaptation and leadership or adaptive leadership assessment scales below demonstrate acceptable levels of CV (face validity, simplicity, relevancy, clarity, necessity), construct validity (convergence and discriminant validity as well as EFA and CFA) (criterion-related validity (concurrence and predictive), and internal consistency/reliability tests (Cronbach test).

- 1.1) Self-efficacy (ALSE)
- 1.2) Shifting reflectiveness (ALSR)
- 1.3) Self-awareness (ALSA)
- 1.4) Adaptive leadership behavior (ALB)

Objective 2: Measuring the Effectiveness of KLC's Leadership Development Path 2023

Applying a longitudinal approach, this objective was defined to address the *teaching and practicing environment gap* and has one main hypothesis.

Hypothesis 2: Participation in KLC's path leads to significant positive change in participants' adaptation and leadership competencies (ALB, ALSA, ALSR, ALSE, and master rationality motive (MRM))

Objective 3: Providing Specific Recommendations for Teaching

This objective was adopted to address the *teaching and practicing environment gap*. It consisted of two quantitative and qualitative phases. The quantitative phase revealed what elements contribute significantly to exercising leadership. Moreover, it provided the teaching team with concrete feedback from the practicing environment through the qualitative phase of the research. Additionally, the qualitative findings were used to explain the quantitative results and provided more details about the involvement of reflective mind components in the adaptive leadership behavior.

Hypothesis 3: There is a positive relationship between adaptive leadership behavior (ALB) and Stanovich's reflective mind components (MRM, ALSR, ALSE, and ALSA) as independent variables (IVs): (Regression model:

$$Y_{ALB} = \beta_0 + \beta_1 X_{SE} + \beta_2 X_{MRM} + \beta_3 X_{SR} + \beta_4 X_{SA} + \epsilon$$

Qualitative Research Question: 1) How do KLC participants describe the decision-making process they take to exercise/not to exercise leadership?

Objectives & Rationale

Objective 1: Developing Instruments for Evaluating LDPs with Adaptive Approach

Under this objective, new leadership scales (ALB, ALSE, ALSA, and ALSR) were developed to evaluate the LDPs employing adaptive leadership approach. The rationale behind this objective was that the KLC's previous assessment tools (e.g., KLC's learning objectives, and particularly Murphy's (1992) leader self-efficacy and Carless et al.'s (2000) leader behavior) had roots in transformational leadership. Hence, they did not match KLC's theoretical framework.

Objective 2: Measuring the Effectiveness of KLC's Leadership Development Path 2023

Under this objective, a longitudinal program evaluation study was designed to evaluate LDPs employing a leadership approach focused on adaptation. The rationale behind this objective was KLC's need to evaluate and improve their LDPs' effectiveness as well as narrowing the training transfer gap.

Objective 3: Providing Specific Recommendations for Teaching

This objective worked toward making teaching modules more targeted based on the participants' involved levels of mind. The connections between participants' reflective components and leadership behavior were studied both quantitatively and qualitatively over a four-month span of time. The rationale for this objective was that identifying the levels of mind involved in exercising leadership helps leadership teachers/coaches target specific mental steps in decision making and build necessary capacity (e.g., cognitive decoupling) as well as reduce Stanovich's (2008, 2011) thinking errors.

Methodological Framework

This research used both quantitative and qualitative methods to evaluate KLC's 2023 path program. Besides program evaluation, the researcher aimed to test an existing theory in cognitive psychology (DPT) by replicating it into leadership field. In fact, this study used both deductive and inductive approach to research.

Quantitative Phase: Deductive-Inductive Research

Per Objective 1, the quantitative phase of this research took deductive and inductive approaches at different points of the scale validation process. Adaptive leadership was used as the framework for developing the scales. Hence, the CV of scale items was deductively tested by a panel of experts. However, since adaptive leadership has not yet solidified into a confirmed

theory, there are no fixed pre-defined theoretical constructs that are anticipated to emerge during the factor analysis. That is why the construct validity of the scale took an inductive approach (EFA) first, followed by a subsequent deductive confirmation step (CFA) (Crawford & Kelder, 2019).

Per Objective 2 and partly objective 3 (Hypothesis 3), the quantitative phase of this research employed a deductive methodology. In deductive research, the study started with a theory and the researcher tried to test the hypotheses derived from it (Locke, 2007; Russel, 2001). The quantitative phase of this research involved two other main theories: KLC's past evaluation results (context level theory) and DPT (as the content level theory) to inform the hypotheses 2 and 3.

The researcher started with this hypothesis, informed by past evaluations, that KLC's programs improves people's leadership behavior. To test this hypothesis, the researcher used *interrupted time-series approach* to rigorously evaluate the path impact. Interrupted time-series design is suitable for the programs in which several data points are available before and after the intervention. According to Newcomer et al. (2015), "Interrupted time-series designs include several observations of an outcome that is expected to be affected by the reform or new program prior to the change and several observations of the same measure after the change" (p. 167). In the research design section, the researcher explained how this longitudinal data collection was implemented.

Besides, informed by Stanovich's (2008, 2011) DPT and empirical findings, the researcher examined how reflective mind is related to adaptive leadership behavior. Hypothesis 3 was designed based on this deductive reasoning. In other words, in this research, an existing

theory was tested by replicating it in the leadership field. DPT informed us about the cognitive processes should be tracked to measure adaptive leadership behavior.

Since Hypothesis 3 was confirmed, the evaluation design was upgraded from interrupted time-series to *regression adjusted covariate design* (Newcomer et al., 2015). In this design, the impact estimates are adjusted using multiple regression by incorporating covariates into the analysis. According to Newcomer et al. (2015), “To the extent that the treated and the untreated groups differ on covariates systematically associated with the outcome, the estimates of program impact can be adjusted for the imbalances between the groups” (p. 167). This systematic evaluation of the variables in a longitudinal manner enhances the rigor of the study. However, since the regression model created was not efficient, the researcher applied the *adjusted covariate design* approach to linear mixed effects model. More details are provided later in this chapter.

Qualitative Phase: Inductive Research

In inductive research, the research does not start with theory. Theory is created by moving from a particular case to general through multiple observations of the phenomenon (Locke, 2007). In the second part of the Objective 3, phenomenology was selected as the methodological theory for two major reasons: its focus on the people’s lived experiences and its emphasis on bracketing. While the former is consistent with the pragmatic ontology of this research and the chosen theoretical framework, the latter becomes crucial in an explanatory mixed-method study, particularly to address potential confirmation bias. How these two become possible by phenomenology is discussed below.

Phenomenology: Lived Experience and Reduction

In phenomenological science, reality is the everyday world we experience. According to Crotty (1998), “the things themselves are phenomena that present themselves immediately to us as a conscious human being” (p. 86). This belief that there are these “things” is what brings the notion of intentionality into attention in phenomenology. Natanson (1973) called it “the axis of phenomenology” (p. 103). Here is where phenomenology introduces its main strategy in following “the back to things themselves” movement in 20th century: Reduction.

Reduction has three steps: epoche, eidetic reduction, and transcendental reduction. Epoche is about bracketing our beliefs, attitudes, and culture. Phenomenology asks us to question our preconceived notions and way of observing the world and just look at what stands before us (Puligandla, 2007). Phenomenologists ask us to suspend our culture and break the inherited understandings of the world. Crotty (1998) says, “Yes. Our culture is liberating. However, ... it is also limiting. It impose[s] meaning, exclude[s] others, and ... harbors its own forms of oppression, manipulation, and other forms of injustice” (p. 87). As such, you need to bracket it while approaching the phenomenon.

Eidetic and transcendental reductions come after epoche. Puligandla (2007) says, “in eidetic reduction one passes from facts to essences. Thus, one may choose anything as the object like a color, figure, relation [and] investigate it from various points of view; for instance, modifications of perception, memory, and fantasy” (p. 20). Then, in transcendental reduction a new consciousness is formed. According to Van Mannen (2014), “[it] is the moment of returning to the world as it shows itself in consciousness” (pp. 91-92). As Sadler (1969) said “phenomenology is an attempt to recover a fresh perception of existence, one unprejudiced by

acculturation” (p. 377). It seems phenomenology has a critical approach that deconstructs and reconstructs reality.

In this study, during the qualitative analysis, the researcher practiced epoche by suspending the quantitative results. For transcendental reduction, the researcher focused on each participants’ transcripts separately and concretely drew out what they described as “fundamental” steps in decision making process using their own words/labels. For eidetic reduction, the researcher moved from descriptions to essences by focusing on the variations in the mental steps people take to exercise leadership. This way, the researcher was able to identify the “essential steps” of decision-making process that were replicable to all participants.

Phenomenological Traditions

There are variety of ways to approach phenomenological methodology. Ethical phenomenology, embodiment phenomenology, hermeneutic phenomenology, Critical phenomenology, sociological phenomenology, political phenomenology, psychological phenomenology (Van Manen, 2014). This work is situated within the sociological-psychological traditions of phenomenological study. Phenomenology is used in a constructivist approach in this research. The researcher, as the human subject, engaged with the individuals who participated in KLC’s training programs through interviews. Focused on their social actions and the feelings around them in their lived and working environment, the researcher tried to make sense of their leadership experiments via affective, process, and causal coding and thematic analysis.

Phenomenological Data Collection and Analysis

It is important to mention that data collection methods like interview in phenomenological inquiry is different from other disciplines like ethnography, biography, or narrative inquiry (Van Manen, 2014). According to Van Manen (2014), “The difference is that

the phenomenological interview and related methods primarily aim to gather pre-reflective experiential accounts” (p. 311). In other words, the researcher dives into participants’ concrete experiences while controlling for their reflection.

Data analysis also has a different approach in phenomenological methodology. The methods of coding and thematizing must be integrated with the method of the suspension and the reduction. Empirical and reflective methods can help the phenomenological analysis. According to Van Manen (2014), empirical methods provide the researcher with experiential material such as actual and imaginal experiences and “reflective methods describe phenomenological reflection in keeping with the reduction proper. In any research project, the selection and usage of empirical (experiential) and reflective (reduction) methods and approaches depend upon the context and the nature of the study” (p. 312). In a mixed-method study like this, the reflective mode of the researcher might be highly influenced by the quantitative results. Therefore, reflection reduction required serious conscious efforts by the researcher.

Phenomenological Interview

This research used interviews as the data collection method using a phenomenological approach, specifically existential phenomenology. According to Schwandt (2007), “this kind of phenomenology is more oriented toward describing the experiences of everyday life as it is internalized in the subjective consciousness of individuals” (p. 226). Phenomenological interview is different from gathering cultural narratives, personal views, perceptions, interpretations, etc. and aims for pre-reflective experiential accounts. Van Manen (2014) declared that “It is much easier for an interviewee to share his or her views, interpretations, or opinions about something than to give a detailed experiential account of an event or moment in a particular place in time” (p. 315). Van Manen (2014) suggested that researchers should try to

capture concrete stories, anecdotes, and experiences of specific situations instead of indulging in speculations and personal opinions. He introduced lived experience description method and recommended asking your interviewees to describe the experience as much as possible as they live through it,

Avoid causal explanations, generalizations, or abstract interpretations. Describe the experience from the inside, as it were almost like a state of mind: the feelings, the mood, the emotions, and so on. Attend to how the body feels, how things smell(ed), how they sound(ed), and so on. Avoid trying to beautify your account with fancy phrases or flowery terminology. (p. 314)

In this study, this way of data collection helped both the researcher and participants implement reduction while doing the analysis and sharing the experience respectively.

Thematic Analysis in Phenomenology

Van Manen (2014) argued that coding and thematizing in phenomenological study is not a mechanical action. “But analyzing thematic meanings of a phenomenon (a lived experience) is a complex and creative process of insightful invention, discovery, and disclosure” (p. 320).

Thematic analysis in phenomenological study is an act of seeing the meaning through a suspension and reduction process. In exploring codes, categories, and themes, we can treat texts as sources of meaning in a wholistic approach (at the level of the whole story); in a selective reading approach (at the level of the separate paragraph); in a detailed reading approach (at the level of the sentence, phrase, expression, or single word) (Van Manen, 2014). For instance, in the detailed approach, Van Manen (2014) asserted that,

We look at every single sentence and ask, “What may this sentence be seen to reveal about the phenomenon or experience being escribed?” Again, we try to identify and

capture thematic expressions, phrases, or narrative paragraphs that increasingly let the phenomenological meaning of the experience show or give itself in the text. If the whole experiential account is particularly powerful, then we try to lift it out as an exemplary story or anecdote. (p. 320)

How this approach was implemented in this research is discussed in detail later in this chapter.

Abductive Research

The first phase of the research takes a mostly deductive approach as it was informed by an existing theory (DPT), adaptive leadership framework, and KLC's past evaluation studies mainly Shams et al. (2023). The second phase of the research employed an inductive approach to research by using phenomenological interview study and bracketing the researcher. However, given the interplay between the two phases of the study and the utilization of qualitative analysis to provide further understanding and interpretation of the results from the initial phase, it could be argued that this study adopted an abductive approach (Tavory & Timmermans, 2014; Timmermans & Tavory, 2012). This became particularly evident when considering the contribution of replicating the DPT in leadership to the existing theory. The following section elaborates how these two phases communicated each other.

Research Design

This project is mixed-method research. The hypotheses were investigated via quantitative methodology while the question was addressed in a qualitative study. Qualitative and quantitative methods were used in a sequential explanatory design. According to Creswell et al. (2003), "The purpose of the sequential explanatory design is typically to use qualitative results to assist in explaining and interpreting the findings of a primarily quantitative study" (p. 178).

Accordingly, in this study, first, the quantitative phase yielded results and surprises about the effectiveness of KLC's LDPs and participants' reasoning processes. Then, the qualitative research helped explain the results and surprises and develop specific recommendations for future KLC paths to turn the results into action.

Objective 1: Developing Instruments for Evaluating LDPs

Quantitative Phase: Scale Validation

Methodology and Procedure. The scale validation came out of an engaged process through which adaptive leadership educators, researchers, and practitioners were involved. Figure 3.1. shows the scale validation process used in this research. First, a list of items measuring leadership ALSA, ALSE, ALSR, and ALB were prepared based on a literature review on adaptive leadership and specifying its domains. Then, these items were shared with a panel of experts (21 members) specializing in adaptive leadership to rate the essentiality (Likert scales 1-3), relevance, clarity, ambiguity, and simplicity (Likert scales 1-4) of items over a month on a Qualtrics survey. Then, item-wise content validity (I-CV) for each individual item and scale-wise content validity for the overall scales (S-CV) were calculated according to formulas below (Boateng et al., 2018; Lynn, 1986; Yusoff, 2019):

$$I-CV = \frac{\text{Number of Experts giving item rating of 3 or 4 for each criterion}}{\text{Total number of experts rated the items}}$$

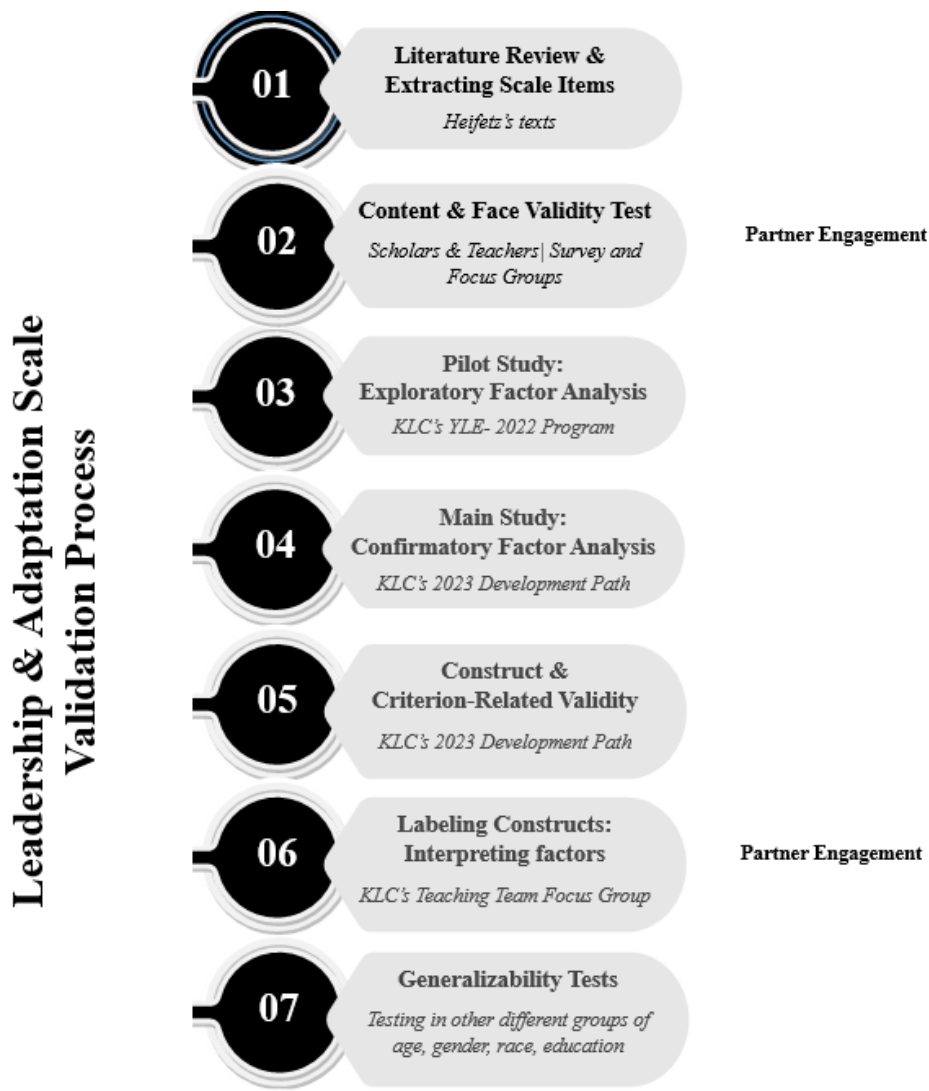
$$S-CV = \frac{\text{Sum of I-CV scores}}{\text{Number of items}}$$

Valid items in terms of content were discussed in terms of face validity in a focus group discussion involving eight members of the panel. KLC's teaching team were participants of this focus group. The focus group discussed two main issues. First, whether to keep or remove the items with minimum CV scores; Second, how to improve the kept items' face validity verbally (clarity, simplicity).

The remaining and modified items were put in a study with a of 249 participants to run EFA to refine the measurement scale items and uncover the latent constructs. After examining the relationship between factors, the hypothesized model underwent CFA using a new set of data collected via surveys asked before each KLC's monthly YLE programs in 2023. It was important to make sure that the CFA data is collected before the program when participants were not exposed to KLC's framework yet. In parallel, for the reliability, internal consistency of the scale items was tested.

The scales were tested in terms of convergence and discriminant validity (construct validity) as well as concurrent and predictive validity (criterion-related validity). Convergent, concurrent, and discriminant validity were tested before KLC's YLE program when people did not receive content about KLC's framework. Finally, the participants' responses were categorized into different racial (URM/ Non-URM), gender, sector, and education groups in order to test the scale in terms of generalizability.

Figure 3.1. Scale Validation Process



Note. The process is adapted from Crawford and Kelder (2019) and Hinkin (1995)

Analysis. Described as below:

Content Validity (CV). Data collected from the panel of experts were tabulated in Microsoft Excel. The I-CVI and S-CVI were calculated for items using the method suggested by Polit and Beck (2006). Items with I-CV relevancy lower than 0.7 or I-CV essentiality less than 0.7 were removed automatically. Generally, items with I-CV relevancy more than 0.7 and essentiality between 0.5- 0.7 were sent to focus group for final decision. Also, items with I-CV

acceptable relevancy (>0.7) and essentiality (>0.5) while one of simplicity, clarity, and ambiguity lower than 0.9 were sent to the focus group discussion for modification discussion. Experts were asked to suggest any corrections they deem necessary to make the items more clear, simple, and certain.

Exploratory Factor Analysis (EFA). Parallel analysis was given priority over eigen value to determine the number of retained factors per Auerswald and Moshagen (2019) instructions. Although the researcher did not solely rely on graphical methods to define the number of factors, Cattell's (1966) scree plot was employed in occasions of *rotational indeterminacy* (Flora & Flake, 2017). Principal Axis as an extraction method was utilized for the datasets that were not normally distributed and for scales with lengthy list of items. Maximum Likelihood was used for datasets that are normally distributed and the scales with shorter list of items.

Rotation methods were also selected meticulously. Orthogonal rotation methods (e.g., varimax, quartimax, and simplimax) were used for scales where factors are assumed to be uncorrelated (e.g., ALSR) while oblique rotation methods (Promax and Oblimin) were used for scales where factors are expected to be correlated (e.g., ALSA, ALSE, and ALB) (Gorsuch, 1983; Osborne, 2019; Tabachnick & Fidell, 2007).

Sampling adequacy for EFA was assessed using Kaiser-Meyer-Olkin (KMO) test (above 0.7) and p-value less than 0.001. To increase the rigor of the process, a more conservative approach than Hatcher (1994) was employed in removing items by keeping the loadings above 0.35. Also, the researcher made sure that the extracted factors' Cronbach's alpha should be >0.70 (Nunnally & Bernstein, 1994).

Minimum sample size for EFA was defined between 200-300 per Guadagnoli and Velicer (1988) and Comrey (1988) suggestion.

Confirmatory Factor Analysis (CFA). After EFA, and generating a hypothesized model, CFA was used to confirm the factor structure using two new datasets with different sample sizes per Boateng et al.'s (2018) guidance. Moreover, in another round of CFA, the datasets were broken into different race, gender, education, and sector categories for which CFA was conducted separately. Addressing Kyriazos (2008) concern that says Chi-square tests and fit indices are sensitive to sample size, using multiple samples, the researcher tried to ensure that the confirmed model is generalizable.

Following Crawford and Kelder's (2019) suggestions regarding the fit indices, the χ^2 , RMSEA (<0.08), Bentler's comparative fit index (CFI >0.9), and the Standardized Root Mean Square Residual (SRMR <0.08) were reported to indicate the model fit. Cronbach's alpha was also assessed to measure the reliability.

Regarding the final decision about the number of scale factors, Lim and Jahng's (2019) recommendation about testing $K-1$ and $K+1$ factors was also considered and tested.

Minimum sample size for CFA was defined 400 per Marsh and Hau's (1999) and Boomsma and Hoogland's (2001) recommendation for models with 2 indicators per factor.

Construct and Criterion-Related Analyses. For these analyses, following Crawford and Kelder (2019) and Hinkin (1995), predictive and concurrent validity (criterion-related) as well as convergent and discriminant validity (construct validity) were tested. Table 3.1. shows the instruments used for testing the construct and criterion-related validity of the scales.

Table 3.1. Criterion-Related and Construct Validity Instruments

Validity	Adaptive Leadership Scale	Instruments
Convergence (Correlation with sister scale)	ALSE	Murphy's (1992) Leader Self-efficacy (MSE)
	ALSR	ALSA
	ALSA	Neider and Schriesheim's (2011) Leader Self-Awareness (NSA)
	ALB	Global Transformational Leadership Behavior (TLB) (Carless et al., 2000)
Concurrent (Correlation)	ALSE	TLB (Same time)
	ALSR	ALB (Same time)
	ALB	Leader Identity (LI) (Hiller, 2005) based on (Bem, 1972; Miscenko et al., 2017)
	ALSA	TLB (Same time)
Discriminant (Correlation)	ALSE	Subjective Job Satisfaction (SJS) (same time)
	ALSR	MRM (same time)
	ALB	SJS
	ALSA	Profession Commitment Scale (PCS) (same time)
Predictive (Predict future-not correlation)	ALSE (T1)	ALB (T3)
	ALSR(T1)	ALB (T3)
	ALB (T1)	LI (T 3)
	ALSA (T1)	ALB (T3)

To interpret the correlation coefficient in concurrent and convergent validity, Ratner's (2009) guideline was used (Table 3.2)

Table 3.2. Pearson Correlation Coefficient Interpretation

Values	Interpretation
0	No linear relationship
+1	Perfect positive relationship
-1	Perfect negative relationship
Between 0 and 0.3/-0.3	Weak positive/negative relationship
Between +0.3 and +0.7	Moderate positive/negative relationship
Between -0.3 and -0.7	Moderate positive/negative relationship
Between +0.3 and +0.7	Strong positive relationship
Between -0.3 and -0.7	Strong negative relationship

Note. Adapted from (Ratner, 2009, p. 139-140)

G*Power (Faul et al., 2007) was used to perform a priori power analysis to detect the correlation between the scales listed above. The effect size for this analysis was estimated conservatively (small to medium) as Cohen's $d=0.3$. The effect size estimate was therefore entered into the power analysis (Correlation: Bivariate Normal Model) with the following input

parameters: Correlation α H1 =0.25, α err prob= 0.05, power=0.95, Correlation ρ H0=0. The power analysis results suggested that N=138.

Additionally, G*Power (Faul et al., 2007) was used to perform a priori power analysis to detect the predictive validity for ALSE, ALSR, ALSA, and ALB. The effect size for this analysis was estimated conservatively (small to medium) as $f^2=0.06$. The effect size estimate was entered into the power analysis (Linear multiple regression: Fixed model, R^2 increase) with the following input parameters: $\alpha=0.05$, power=0.80, number of predictors=5 (including controls: Race, education, gender, age). The power analysis results suggested that N=133.

Internal Consistency Test. Cronbach alpha was used to test the internal consistency of the items within each scale, ensuring the minimum threshold of 0.7.

Generalizability Test. To make sure that the instruments are replicable, the developed model underwent testing in different demographic groups and time (race, education, gender, and time) following Cronbach et al.'s (1963) generalization theory. However, achieving the perfect generalizability based on this theory requires addressing other aspects as well, such as different LDP and settings, which suggest the trajectory of the researcher's future studies.

Qualitative Phase: Finalizing Items and Scale Factor Labeling

Methodology and Procedure. This involved two focus group discussions for making final decision about keeping or removing scale items after CV test and labeling factors after CFA. Members of the panel of experts were invited via email to the first focus group to discuss the face validity of items and decide whether to keep or remove the items with minimum I-CV. This 60-minute online discussion was facilitated by the researcher on Zoom. See the Agenda in Appendix J.

Focus Group 1 (Finalizing List of Items). An online 60-minute focus group discussion was designed after the CV test (see the session design in Appendix J). Members of the panel of experts were invited to this discussion via email.

First, the researcher as the facilitator presented a short report on the CV results. The items that were automatically removed and items that needed to be discussed both in terms of content and face validity were shown to the group. Then, the facilitator formed two small groups and assigned each group a set of 13 items (items that needed a final decision) to discuss whether they wanted to keep the items or remove them. After the small group discussions, each group presented the list of items they decided to keep accompanied by their arguments while the other group was supposed to critically ask questions.

In the third section of the focus group, the participants were asked to discuss items that needed wording revision to improve in terms of face validity (simplicity, clarity, and ambiguity). The researcher showed items one by one on a single slide and presented the suggestions that the panel of experts made in the CV survey. The group was asked to discuss potential revisions and reach a consensus on the wording of each scale item.

Focus Group 2 (Scale Factor Labeling). Another focus group was taken place online on Zoom one year later after running the EFA and CFA tests to label the scale factors. Again, members of the panel of experts particularly KLC team as the partner were invited to this 90-minute session (Appendix J). First, the researcher presented a review of the scale validation process and defined the word “Factor”. Then, participants were invited to two small groups with a minimum of three participants for 30 minutes. They were asked to give a maximum of two unique labels to each factor covering the content of all assigned items. The researcher clarified that each label should be unique enough to make its items distinct from other factors and broad

enough to encompass the content of all the included items. Moreover, they were encouraged to use the concepts of adaptive leadership framework for labeling. During this 30-minute small group discussion, the researcher actively circulated between the two groups, providing clarifications if needed. After each group finalized their labels, the researcher facilitated a large room discussion to reach a consensus on the final label for each factor.

It should be noted that this phase was conducted asynchronously for ALSR due to shortage of time in the online session.

Analysis. No analysis was needed on the focus group discussions. The researcher's goal for these meetings was to reach consensus. Both focus group discussions continued until consensus was reached on the final decision about removing items, improving items' face validity, and labeling each scale factor.

Objectives 2 & 3: Measuring Effectiveness | Recommendations for Teaching

Quantitative Phase

Methodology and Procedure. After validating the scales, they were used for evaluating KLC's path in 2023 using an interrupted time-series evaluation design (Newcomer et al., 2015). The leadership development path consists of three training programs that two of them were studied in this research: YLE and LFC. The participants received KLC's teaching modules for 45 hours (A two-day and a three-day program over one year). The study used the pre-post surveys on Qualtrics in alignment with the first and third level of Seemiller and Rosch's (2019) six-domain model of leadership development (level 1: efficacy and motivation- level 3: performance). It evaluated the participants' ALSE, ALSR, MRM, ALSA, and ALB at four timepoints: before starting YLE/LFC (T1), right after finishing YLE (T2), one month after YLE (T3), four months after YLE (T4)/three months after LFC (T2). The variables were compared

across the timepoints within and between the groups. All validated scales are on a seven-point Likert scale, Strongly Disagree (1) to Strongly Agree (7) except ALSE (0-100). The ALB survey was not asked at T2 as change in performance was not expected at the end of the program.

Analysis. To test Hypothesis 2, first, participants' data were matched across timepoints. Incomplete cases above the threshold of three missing items (for ALB) and one (for other scales) as well as the responses with a duration less than half of the time recommended by Qualtrics were removed. Responses with Z score beyond (-3, +3) were considered as outliers and were removed from the dataset. The researcher ensured the acceptable internal consistency indicated by Cronbach's alpha scores of at least 0.70. It was confirmed the data do not violate the assumption of the normality of the different scores. In cases of violation, the data was transformed and corrected to be normally distributed using Box-Cox or SQRT transformations. Generally, in this research, normality is confirmed if histograms of the difference scores appear normal and skew values < 3 and kurtosis <10. Linear mixed effects models (LMM) were run to measure the effect of Time (program) on each variable.

To test Hypothesis 3, G*Power (Faul et al., 2007) was used to perform a priori power analysis to detect the relationship between ALB, ALSE, ALSR, ALSA and MRM and also other dummy variables such as gender, race, and education. The effect size for this analysis was estimated conservatively (small to medium) as $f^2=0.09$. The effect size estimate was entered into the power analysis (Linear multiple regression: Fixed model, R^2 increase) with the following input parameters: $\alpha=0.05$, power=0.80, number of tested predictors=4, total number of predictors=7. The power analysis results suggested that N=138. We chose this statistical test since we wanted to see how adding predictor variables will increase the explanatory power of the regression model.

Additional Analysis. Since the regression model did not converge successfully, LMMs were used to come up with a more precise equation. According to Meteyard and Davies (2020), “LMMs are warranted when data are collected according to a multi-stage sampling or repeated measures design. That is, when there are likely to be correlations across the conditions of an experiment because the conditions include the same participants...” (p. 1). Besides the inefficiency of the hypothesized regression model, LMM was preferred over regression model in this research for four other reasons (Feehley et al., 2019; Murphy et al., 2022; Wolfinger et al., 1991; Rosenkranz, 2015):

- 1) Based on our theoretical framework and repeated measure design, we assume that the coefficient of predictor variables changes across time.
- 2) Based on the theoretical framework, adaptive leadership behavior cannot be captured merely through studying reflective mind processes. LMM was used to recognize other random effects which account for differences between subjects (potentially algorithmic mind cognitive processes). LMM helped with understanding what percentage of adaptive leadership behavior can be attributed to fixed and random effect of predictor variables as well as random effects of intercept which accounts for individual differences.
- 3) Since participants responded to the same questions across three timepoints, to increase the accuracy of the model by correct modelling of longitudinal measurements correlation LMM was used
- 4) Due to lack of seamless data collection system, there are several missing data across measurement times. LMM helped with handling the missing data through maximum likelihood estimation.

Four different scenarios were run – null, basic, intermediate, and advanced. We started with a null model accounting for random effects on intercepts. For the basic level, the predictor variables were isolated in the model. Regarding the intermediate scenario, predictor variables were added in interaction terms. In the advanced scenario, random slopes were loaded on the intermediate models. Model choice happened through best fit comparisons following an integrated approach of considering Likelihood Ratio Tests (LRTs), Akaike Information Criterion and Bayesian Information Criterion (AIC/BIC), and minimal model approaches simplifying model by removing non-significant fixed or random effects (Meteyard & Davies, 2020). For this type of analysis, literature says, minimum 50 sampling units per measurement is needed to run LMM (Bell et al., 2010; Maas & Hox, 2004).

Qualitative Phase

Methodology and Procedure. Objective 3 was partly addressed by the qualitative phase of this mixed-method study. YLE's T4 survey included a question asking if the respondents would be willing to be interviewed (two times) about their leadership experiences. KLC provided the researcher with the survey database and the names and emails of respondents. The researcher contacted those participants who completed all surveys and agreed to have the interview via email to schedule dates and times on Calendly.

In the existential phenomenological interviews, the interviewer tried to capture concrete stories, anecdotes, and experiences of specific situations instead of indulging in speculations and personal opinions (Van Manen, 2014). In these semi-structured interviews, participants were asked to describe their lived experiences as much as possible. Processes of action (especially dynamics of time), emotions, and motivations would give the researcher a precise frame of focus both in interviewing and analyzing the data.

Round 1 Interview Questions: Building Trust/Rapport and Focus on Successful

Experiments (30-45 mins). Interviews started with five general and opening questions:

- How do you describe your leadership role in your work/community?
- What are your aspirations for your organization/community?
- How were the training programs? What went well and what could have been better?
- How do you conceive leadership today?
- If you go back, what would you like to see differently in this program?

At the beginning of the interview, the researcher explained that this interview is not a quiz or interrogation to know whether they practiced what they were taught or not. It was clarified that this interview does not aim to know whether the participants remember or learned what that were taught in the path. On some specific occasions during the interviews, the researcher emphasized that the main purpose of asking straightforward follow up questions was to delve deeper into their thought processes and leadership experiments, whether successful or failed.

In this round, concrete stories of participants' successful experiments were delved into via asking below questions:

- a. Can you walk me through a recent time in which you exercised leadership successfully?
- b. As the follow-up questions: Can you describe the specific moment that you decided to do that? What was the challenge? What came to your mind? How did you analyze the situation? How did you get motivated? What was your thought process before, during, and after intervention? Can you describe the environment around you in that moment? Do you remember anybody or anything that helped you? How did you feel?

How did you manage that emotion? How do you feel now about your feelings at that moment? Did you experience that in other occasions? Can you walk me through that? (Pugh, 2013)

- c. How do you know that your intervention was successful?

Round 2 Interview Questions: Failed Experiments (1 hour). In this round, the researcher focused on the concrete stories of participants failed and/or non-action experiments by asking below questions:

- a. Can you please walk me through a recent experience in which you think there was room for improvement in your action?
- b. What do you think impeded you from doing that successfully? Can you describe that moment? How did you analyze the situation? What were some thoughts that came to your mind holding you back from action? What was your thought process before, during, or after intervention? Can you describe the environment and the people around you? How were their reactions? How did you feel? How did you manage those emotions? How do you feel about your feelings in that moment? How do you feel now about that moment? What happened after that intervention? Did you experience that in other occasions? Can you walk me through that?

Interviews were started with six participants and a preliminary phase of analysis initiated soon after. Then, the researcher embarked on a second phase of interviews with four more participants to reach the saturation in data analysis. The researcher continued the interviews until no disconfirming evidence was found.

Analysis. In this phase of the research, phenomenology was used as a qualitative methodology to corroborate (to some extent) and explain the quantitative results. The inductive process as the analytic method and NVivo as the management software were employed.

First Cycle Coding. The mental steps people took to finally decide to exercise leadership were explored. The researcher started with first cycle coding, for each participant separately, with a blend of coding methods in which initial codes were derived: process, emotion, motivation/value, and causal codings. Since leadership here is an interactional activity, process coding could help with identifying those reasoning steps (Saldaña, 2021). According to Saldaña (2021), “processes imply actions intertwined with the dynamics of time, such as those things that emerge, change, occur in particular sequences” (p. 142). Process coding could also show how different types of behaviors might take different ways of reasoning. Process maps were drawn for each participant separately encompassing all the details such as emotions and motivations described. The researcher went through this initial cycle as a strategy for transcendental reduction.

Participants were asked to share their motivations (values, beliefs, etc. as dispositions) as well as describe their feelings before, during, and after the intervention moment. Therefore, the researcher had a specific focus on emotions and motivations in the decision-making process. Emotions and values can be coded. According to Saldaña (2021), “emotion coding tap into the inner cognitive systems of participants and values coding assesses a participants’ integrated values, attitudes, and belief systems at work” (p. 156). In this research, theoretically, emotions and values are considered as triggers of rational action that can help identify specific cognitive processes in decision making. As such, it was assumed that emotional and value coding are suitable tools for identifying participants’ beliefs, values, and attitudes as well as tracking

emotional journeys before, during, and after a leadership action (Saldaña, 2021). These three types of coding are aligned with the selected phenomenological methodology.

Here are questions that the researcher asked himself while coding each paragraph of the interview transcripts:

- What are people doing? What are they trying to accomplish?
- How, exactly, do they do this?
- What specific means and/or strategies do they use?
- What is the reason behind this?
- What do I see going on here?
- What did I learn from these notes? Why did I include them?
- How is what is going on here similar to, or different from, other incidents or events recorded elsewhere in the other notes?
- What are the surprises, anomalies, and disturbs here?
- What are the feelings and emotions here?
- What honorable (admirable behaviors and acts), schematic (the jokes, metaphors, etc.), the visceral (facial expressions, pauses, silences, laughter, etc.) information are available here?
- What meta-feelings can be seen here? (How do they feel about how they feel?)

In addition to actual coding, pre-coding was used during the interviews and analytic memos accompanied with. Reading each transcript twice and focusing on the thought processes involved in decision making to act, a unique process for each participant's actions (success and failed) was coded on NVivo and mapped on whiteboards besides memo notes. Hence, 20 unique memos and process maps were drawn overall for 10 participants in the first cycle of coding.

Second Cycle Coding. Involves “classifying, prioritizing, integrating, synthesizing, abstracting, conceptualizing, and theory building” (Saldaña, 2021, p. 89). Transitioning from first to second cycle coding requires constructing concepts from categories, reorganizing and reassembling the data for better focus on the direction of the study, and drawing preliminary models of the primary actions (Saldaña, 2021). Saldaña (2021) describes this as axial coding. “This method aims to link categories with subcategories and asks how they are related and specifies the properties and dimensions of a category” (Saldaña, 2021, p. 307). In this research, concept mapping was used to do so. Selecting the main codes and removing the redundant ones, the reasoning steps and their associated dimensions were categorized and subcategorized by asking “if, when, how, why” questions from what happened and not happened. Through second cycle coding, the researcher moved from facts to essences; from transcendental reduction to eidetic reduction.

In transitioning to second cycle coding, initially, the 12 unique process maps of the first six participants were arranged next to each other on different boards. Then, the researcher started to merge the process maps into one integrated diagram using Office word (The researcher found NVivo limited in designing an integrated concept map). The diagram was created as the skeleton of the final concept map not only to display analytic categories but also to provide the answers to the research questions. Then, the first draft was modified after finishing the second phase of interviews and analysis taking the same analysis process as the first phase.

After finalizing the codes, subcategories, categories, and themes and creating final integrated concept map, the researcher initiated theming to theorizing process. It was expected that final themes or categories would lead to a theory that reveals the reasoning process toward exercising leadership. According to Saldaña (2021), this theory “expresses a patterned

relationship between two or more concepts; predicts and controls action through if–then logic; explains how and/or why [leadership] happens by stating its cause(s)” (p. 350). This can help the researcher with providing specific recommendations for future LDPs.

Analytic Memos and Researcher Reflections. Analytic memos are the researcher’s notes to self about the “participants, phenomenon, or process under investigation by thinking and thus writing and thus thinking even more about them” (Saldaña, 2021, p. 57). Memos were helpful in drawing the decision-making process which involved emotions, values, and changes in behaviors. While coding the raw data, analytic memos about the following topics were prepared:

- Participants’ actions, reactions, emotions
- What I find intriguing, surprising, and disturbing
- Emergent patterns, categories, themes, and concepts
- Emergent theory
- My own critical dilemmas
- Future directions for the study
- Meta memos
- Tentative answers to research questions.
- Passage drafts for the final report
- Biases and subjectivities.

Also, memos helped with transitioning from first to second cycle coding. Moreover, these memos helped later find the alignment between the findings in the qualitative research with the quantitative results as well as the content level theory. During coding, process mapping, concept mapping, the researcher experienced several occasions when the content level theories (DPT and CDT) directing him toward certain codes and categories. Therefore, to bracket himself, he used

Van Manen's (2014) four themes of analysis (lived time, lived body, lived space/things, and lived relations) to distract himself from preconceived codes. However, finding it impossible, frequent reflection moments were attended intentionally to assess whether there is a balance between content level theory directions and what data supports. To be conscious of this in the analysis and writing, the differences and surprises were highlighted in the memos.

Mixed-Method Analysis

After separate analyses, the researcher tried to compare and contrast quantitative and qualitative results. In cases of correspondence, first, coherence of the two analyses were highlighted and then quantitative findings were explained by qualitative results. Saldaña (2021) called this paradigmatic corroboration and defines it as what "occurs when the quantitative results of a dataset do not simply harmonize or complement the qualitative analysis but corroborate it" (p. 38). In case of conflicts or surprises, a new path opened for interviews to focus on the disconfirming evidence or future studies were recommended.

Data Representation

Quantitative Phase

Quantitative data collected, techniques, statistical procedures, and anticipated and unanticipated results were presented in diagrams, charts, and tables as well as texts (See Appendices). The results related to each hypothesis were discussed. Results that needed to be explained or explored further in the qualitative research were underscored.

Qualitative Phase

After presenting the final theme as the theory and the concept map, arguments were organized in paragraphs having categories and subcategories as headings and subheadings. Codeweaving technique was employed to write the paragraphs. Saldaña (2021) defined

codeweaving as “the actual integration of keywords from coding into the analytic memo narrative; a practical way to ensure that you are thinking how the puzzle pieces fit together” (p. 65). In paragraphs, codes, categories, themes, and concepts of the analysis were woven into sentences to demonstrate the theory. Direct quotes were used to bring more illustrative and demonstrative evidence. Pseudonyms and positions of the participants were used in the qualitative descriptions to add a humanizing touch to the text. After preparing the first draft, the report was sent to interviewees for member check for one month.

Validity & Reliability of the Research

Quantitative Research

Validity and reliability both show the efficiency of a scale/measurement as an assessment tool. Validity refers to the accuracy of a measure/scale while reliability refers to the consistency of the measure. In other words, validity responds to this question: “Does the scale/tool measure what it is supposed to measure?” whereas reliability responds to this question: “Does the scale/tool reproduce consistent results over time?” Table 3.3. shows how the researcher ensured validity and reliability of the quantitative results. Details were discussed in the research design.

Furthermore, it is important to note that during the CV test, the researcher adopted a conservative approach when deciding whether to remove items. Despite the CV results suggesting the removal of several items, the researcher opted to take an additional qualitative step. This involved subjecting some of the items to focus group discussions to make the final decision.

Qualitative Research

Validity and reliability in qualitative research cannot be defined based on quantitative standards. Hence, evaluating the quality of qualitative research requires distinct strategies that

might not be applicable to quantitative research. This distinction in qualitative research's validity and reliability standards is more clarified below.

Validity

Validity in qualitative research is measured by rigor and face validity. Readers can assess the validity of qualitative research based on its rigor. According to Tracy (2010), in interviews, demonstration of rigor includes “number and length of interviews, appropriateness of the interviews sample, types of questions asked, the level of transcription detail, and the practices taken to ensure transcript accuracy” (p. 840). In this research, the interview guideline, number of interviews, and transcripts were brought as evidence that can help readers assess the rigor.

Also, qualitative research can demonstrate validity by affecting the audience (Resonance). According to Tracy (2010), “The potential of research to transform the emotional dispositions of people and promote greater mutual regard has been termed ‘empathic validity’” (p. 844). Readers can assess the validity of this research by its resonance which emerges out of the potential transferability of mental processes described across different leadership development programs. The researcher tried to achieve this transferability through an evocative description (i.e., contains feelings and emotions) of the decision-making process that people take to exercise leadership. It was expected that, by reading the description of the distance between learning and exercising leadership in this study, the readers would intuitively transfer the results to their own leadership experiences (Tracy, 2010). This contributes to the research replicability as well.

Reliability

Unlike quantitative research, qualitative researchers look for surprises. The more the researcher looks for negative cases, the more reality appears. This was the primary rationale behind the researcher's decision to add a subsequent qualitative phase to the research. As such, it

seems reliability in qualitative research depends on how accurate, precise, and broad the phenomenon is presented (Becker, 2001) not only through analysis but through what the data have shown. That is why Katz (2001) said, “to insulate a careless analysis from critical readers, the researcher would have to engage in a laborious process” (p. 142). As such, readers can assess the reliability of this research by attending Tracy’s (2010) credibility criterion. The researcher invites readers to assess this research based on the thick description it provides of the reasoning processes people described. Also, readers can assess the credibility of this research by observing how the results of the quantitative and qualitative phases and explanations of content level theories converged on the same conclusion (Triangulation).

Table 3.3. Validity and Reliability Tests Used

Hypotheses/ Question	Validity/Reliability		Tests
H1	Validity	CV	Essentiality
			Relevancy
			Clarity, Simplicity, Ambiguity
		Criterion -Related Validity	Face validity test
			Concurrent
			Predictive
			Convergence
		Construct Validity	Discriminate
			Generalizability
	Reliability	EFA	Sample adequacy test: KMO
			Cattell’s Scree Plot
			Loading power
			Cross loading
			P-Value
			Cronbach α
		CFA	χ^2
			RMSEA
			SRMR
			CFI & TLI
H2-3	Validity & Reliability	LMMs	Cronbach α
			AIC
			BIC
			LogLikel.
			F-tests
			LRT
			Pearson Correlation Coefficient
			R-squared marginal and conditional
			Kendall Correlation
Q	Validity		Intraclass Correlation Coefficient (ICC)
			Empathic validity
			Face Validity
	Reliability		Resonance
			Rich Rigor
			Triangulation
			Credibility (Saturation)
			Multivocality
Member check			

Engagement Strategy

The scale development and program evaluation in this research followed an engaged scholarship approach. Third Floor Research as an institutional connector between university and KLC provided a unique opportunity to bridge the knowledge of academia and practice of the LDPs provider. Affiliated with this applied learning initiative as a graduate research assistant for three years, the researcher followed Bryson and Patton's (2015) instructions for engaging stakeholders in developing assessment tools and interpreting the evaluation results:

- 1- *Developing facilitation skills:* During the PhD program, the researcher took courses and benefited from the available opportunities to facilitate discussion sessions. He also facilitated several discussion sessions with the program providers for other research studies as part of his responsibilities as a research assistant. This way researcher built minimum competence and confidence needed for facilitating discussions in this research.
- 2- *Find evaluation information users:* This was achieved by attending KLC's programs and connecting to their participants, teachers, and staff members during the three-year collaboration as a research assistant.
- 3- *Find tipping point connectors:* Collaborating with KLC through Third Floor Research and working directly with KLC's director of research, the researcher established connections with KLC's CEO, chief impact officer, chief civic leadership development officer, program director, and director of communications. These key individuals played pivotal roles as connectors, bridging research findings with teaching practices in this study.

- 4- *Facilitate high quality interactions:* The researcher, in collaboration with KLC's director of research (major advisor of this dissertation), endeavored to enhance the structure and purpose of interactions by establishing a focus group comprising tipping point connectors and evaluation users. They conducted group discussions at various stages of assessment tool development, including content validity testing, face validity testing, and interpretation of factors.
- 5- *Nurture interest in evaluation:* This has been a core strategy in Third Floor Research. One of the main activities of Third Floor Research and the researcher has been to show the theoretical/empirical foundations of KLC's framework and explain why their teaching practice is impactful. As an internal strategy, this was addressed through frequent presentations and formal and informal conversations with KLC's teaching team and staff, who served as evaluation users and tipping point connectors in this research.
- 6- *Anticipate turnover of the intended users and participants' dropout:* Anticipating missing data, incomplete survey responses, and non-responses to longitudinal surveys, the research team took proactive measures from the outset based on past experiences. These included offering incentives for post-program data collection, maintaining continuous engagement with participants, and engaging KLC's director of programs and teachers in encouraging participants and communicating the importance of data collection with them.

Ethics

The research proposal was submitted to IRB and received approval (Appendix Z). Specific consent forms were designed for quantitative and qualitative phases of the research including surveys, focus groups, and interviews. It was mentioned, in IRB, that all the participants' responses would be kept strictly confidential. The researchers may publish the results of the study. If they do, they will only discuss the aggregate results of all interviews. Moreover, before publishing the qualitative part of the research, the participants received a zero draft of a research summary to voluntarily read and give their final comments. Their names were replaced with pseudonyms in this volume and will not be used in any publication or presentation about the study. In terms of reciprocity and mutual benefits, it was mentioned in the consent form that this research might not have direct benefits for the participants. However, the results and final report will be published by KLC and shared with participants and partner organizations. The report can inform the organizations about possible barriers in their processes and structures toward leadership actions.

A research dilemma was the disclosure of interviewees' information: internal and external confidentiality (Tolich, 2004). One might ask, is that possible that KLC discloses information about the internal environment of an organization to the public or court? The response is that the name of organizations will be kept in secret or changed to pseudonyms if necessary. Besides, as KLC believes in process of leadership progress in organizations, internal challenges mentioned in the interview will not be disclosed to the public. Reviewing KLC's previous reports, it is possible to disclose those challenges in future reports upon resolution by the organization as the evidence of progress (Ng, 2021). This usually happens over years in books-e.g. *When everyone leads*- or stories published in KLC's monthly journal.

The research consent forms and IRB application needed modifications during the research. For instance, due to not having enough response rate on T4 data, the researcher decided to provide incentives using KLC's financial support. Moreover, the researcher changed the predictive validity and discriminant validity test scales following the preliminary analysis on scale validation results.

Highlights of the Chapter

- A methodological framework in alignment with the theoretical framework (presented in Chapter 2) was developed.
- A rationale for conducting a mixed-method study with a sequential explanatory design was presented.
- Rationale for each objective was explained.
- Research design was presented in detail comprising of quantitative and qualitative phases.
- Details of quantitative data collection and analysis standards for scale validation process (CV, EFA, CFA, generalizability tests) were presented.
- Details of qualitative data collection (focus group and phenomenological interviews) and analysis process were described.
- Distinctions were made between the methods used to assess validity and reliability in the quantitative and qualitative phases.

Chapter 4 - Findings & Discussion

This chapter presents a detailed description of how the methodology described in the previous chapter was implemented in relation to the theoretical framework. As mentioned earlier, this research involved eight main rounds of data collection with different methods:

- 1) Content validity test (CV) (Survey)
- 2) Face validity test (Focus group discussion)
- 3) Exploratory Factor Analysis (EFA) (Survey)
- 4) Confirmatory Factor Analysis (Survey)
- 5) Labeling constructs (Focus group discussion)
- 6) Your Leadership Edge (YLE) program evaluation
- 7) Leader For Change (LFC) program evaluation
- 8) Qualitative interviews

All these phases of data collection and analysis are connected. While they worked toward their specific objectives introduced in the first chapter, they altogether served the main aim of this research which is the development of theory-grounded valid, and reliable assessment tools and model for LDPs with an adaptive leadership framework. In this chapter, first, the demographic characteristics of the different pools of participants for each round of data collection and analysis were described. Then, findings were presented in response to the quantitative hypotheses and qualitative questions one by one. Finally, an in-depth analytical discussion was provided to make sense of the findings and theoretical and practical connections were built between qualitative and quantitative results.

Demographics

Content Validity Demographics

After extracting items from the literature, the list of items (for all four scales) was emailed directly to the panel of experts in the format of a survey on Qualtrics. Within a month and after sending two reminders, 21 professionals with a background in teaching, coaching, research, and practicing adaptive leadership responded to Relevancy, Clarity, Simplicity, and Ambiguity questions while 11 experts responded to Essentiality questions since it was added to the survey with a one-week delay. KLC's teaching and coaching team as the partner and main beneficiary of this research constituted nearly half of the respondents. Nine faculty members practitioners, and researchers (out of 21) outside KLC also responded to the survey. Table 4.1 shows the panel of experts' background and connection to adaptive leadership framework.

Table 4.1. Demographic Characteristics of the Panel of Experts

Experts	Background	Connection to Adaptive Leadership Framework
Expert 1	Leadership Studies (MSc.)	Instructor-Researcher-Practitioner
Expert 2	American Studies (B.A)	Teacher-Coach-Practitioner
Expert 3	Educational Leadership (MSc.)	Teacher
Expert 4	Marketing (B.A)	Practitioner
Expert 5	Communication	Teacher
Expert 6	Organizational Psychology	Instructor-Researcher-Practitioner
Expert 7	Organizational Psychology	Researcher
Expert 9	Leadership Studies (PhD)	Instructor-Researcher-Practitioner
Expert 10	Leadership Studies (MSc)	Instructor-Practitioner
Expert 11	Educational Leadership (PhD)	Instructor
Expert 13	Communication (PhD)	Instructor-Researcher-Practitioner
Expert 14	Leadership & Dramatic Art	Teacher-Coach-Practitioner
Expert 15	Organizational Management & Leadership	Teacher-Practitioner
Expert 16	Leadership Studies (PhD)	Instructor-Researcher
Expert 17	Psychology (PhD)	Researcher-Coach-Teacher
Expert 18	Leadership (PhD)	Teacher-Coach
Expert 19	History (B.A)	Teacher-Coach
Expert 20	-	Teacher-Coach
Expert 21	Language & History (M.A)	Practitioner
Expert 22	Political Science	Teacher-Coach
Expert 23	Educational Leadership (Ed.D.)	Instructor

EFA Demographics

After refining and editing the list of items based on the content validity test results (discussed later in the chapter), the items were put in a separate study to run the EFA. 265 participants responded to this survey from different pools: KLC's YLE-November 2022 (42.1%) and YLE- December 2023 (36.5%), K-State Leadership Minor Students (2.6%), Cleveland Leadership Center's (CLC) September participants (5%), and the researcher's network (18.8%). The average age of the respondents was 42.3, with the youngest being 19 and the eldest was 69. Most participants had a bachelor's or master's degree (48.7%) and identified as White or Asian (Non-URM) (60%), while underrepresented minorities (URM) constituted 9.8%. Participants were at different career stages and worked in various sectors (i.e., governmental, private, non-profit, education, and faith) (Table 4.2).

Table 4.2. Demographic Characteristics of the EFA Dataset

Baseline Characteristics	N=265	%
Cohort		
Pool1 – YLE November 2022	112	42.3
Pool 2 – YLE December 2023	97	36.6
K-State Minor Pool – April 2023	7	3.6
Personal Networks – May 2023	37	14
CLC LDP September 2023 Participants	12	4.5
Level of Education		
High School	27	10.2
Some College	11	4.2
Associate's Degree	16	6
Bachelor's Degree	68	25.7
Master's Degree	61	23
Doctorate Degree	16	6
Missing/Preferred not to respond	66	24.9
Race		
URM	26	9.8
Non-URM	159	60
Missing/Preferred not to respond	80	30.2

CFA Demographics

After EFA, a new survey with the shortened list of items was sent to a larger group of participants to collect data for the CFA. Participants were selected from KLC's YLE 2023 cohort (Except December) and contacted right before starting the KLC's leadership path. The researcher made sure that nobody responded to this survey after starting the program to ensure the reliability of the CFA analysis. To increase the generalizability, the CFA was run on two datasets per Boateng et al.'s (2018) suggestion. Dataset 1 is a pool of 350 participants and Dataset 2 has 603 participants. These two datasets were separated from each other based on the duration of data collection (Each pool includes responses from about five months). The next two sections describe the demographic characteristics of these two pools. Then, the two datasets were merged and broken down into different categories based on gender, race, education, age group, time, and sector to run another round of CFAs separately. The demographics of these categories were not presented to avoid redundancy.

Dataset 1

This dataset contains 350 responses from the first five cohorts of YLE 2023 program. The April cohort had the largest response rate while the June cohort the smallest. The average age for this pool was 41.8, with the youngest being 21 and the eldest was 78. Most participants had a bachelor's or master's degree (63.8%) and identified as White or Asian (Non-URM) (74.9%), while underrepresented minorities (URM) constituted 14.3%. Most of the participants were female (69.4%). Participants were at various career stages and worked in various sectors (i.e., governmental, private, non-profit, education, and faith). There were several participants who preferred not to respond to demographic questions particularly education level, gender, and race. Participant demographics of this pool can be found in Table 4.3.

Table 4.3. Demographic Characteristics of the CFA -Dataset 1

Baseline Characteristics	N=350	%
Cohort		
Cohort 1 – Feb 2023	50	14.3
Cohort 2 – March 2023	88	25.1
Cohort 3 – April 2023	99	28.3
Cohort 4 – May 2023	85	24.3
Cohort 5 – June 2023	8	2.3
Missing	20	5.7
Level of Education		
High School	14	4
Some College	27	7.7
Associate’s Degree	29	8.2
Bachelor’s Degree	108	31
Master’s Degree	115	32.8
Doctorate Degree	14	4
Missing/Preferred not to respond	43	12.3
Race		
URM	50	14.3
Non-URM	262	74.9
Missing/Preferred not to respond	38	10.9
Gender		
Man	69	19.7
Woman	243	69.4
Non-Binary	2	0.5
Missing/Preferred not to respond	36	10.3

Dataset 2

This dataset has 603 participants who attended the YLE program within the second half of 2023. This pool exhibited a diverse range of demographics, with an average age of 42.8. The age spectrum spanned from 13 to 72, showcasing a broad representation. A significant portion of participants held bachelor’s or master’s degrees, accounting for 47.4% of the group. The majority identified as non-URM (69%), while URM participants formed 13.3%. Gender distribution was notable, with 61.4% of participants identifying as female. The participants

displayed varied career stages and professional backgrounds, engaging in five main sectors such as governmental, private, non-profit, education, and faith-based organizations. See Table 4.4 for a detailed breakdown of participant demographics.

Table 4.4. Demographic Characteristics of the CFA- Dataset 2

Baseline Characteristics	N=603	%
Cohort		
Cohort 5 – June 2023	70	11.6
Cohort 6 – July 2023	79	13.1
Cohort 7 – August 2023	132	21.9
Cohort 8 – September 2023	102	17
Cohort 9 – October 2023	91	15.1
Cohort 10 – November 2023	43	7
Cohort 11 – December 2023	33	5.5
Missing	53	8.8
Level of Education		
High School	22	3.6
Some College	38	6.3
Associate’s Degree	43	7.1
Bachelor’s Degree	154	25.5
Master’s Degree	132	21.9
Doctorate Degree	19	3.2
Missing/Preferred not to respond	195	32.3
Race		
URM	80	13.3
Non-URM	416	69
Missing/Preferred not to respond	107	17.7
Gender		
Man	130	21.6
Woman	370	61.4
Non-Binary	1	0.1
Missing/Preferred not to respond	102	16.9

YLE Program Evaluation Demographics

YLE Total Participants

1634 participants attended YLE 2023 monthly programs. This is the same population that the CFA datasets were extracted from. 46.2% of them participated in the in-person programs in April, May, July, September, and October while 53.8% of them the virtual version in March, February, June, August, November, and December. The average age of YLE participants was 42.5, with the youngest being a 13-year-old student and the eldest being 78. Most participants had a bachelor's or master's degree (53.9%) and identified as White or Asian (Non-URM) (73.4%), while underrepresented minorities (URM) formed 16.3%. Most of the participants were female (66%). Participant demographics can be found in Table 4.5.

YLE Sample Participants

Two different datasets were used to run the program evaluation analysis. Dataset 1 comprising 141 participants who responded to all T1, T2, and T3 surveys was targeted to trace the YLE impact on ALSR, ALSA, MRM, and ALSE. Dataset 2, with 160 participants who completed surveys at T1, T3, and 84 participants completed T4, was specifically used to evaluate YLE's impact on ALB.

Due to the overlap between the two groups, demographic information for Dataset 2 is the only one presented here. The average age for this pool was 44.8, with the youngest being 22 and the eldest 74. As Table 4.6 shows, the demographic characteristics especially gender, education, and race breakdowns, closely resemble those of the YLE population. This validates the suitability of this sample for conducting the analysis. It should be noted that December cohort data was used for program evaluation, not for the CFA, as it was used for EFA complementary phase.

Table 4.5. Demographic Characteristics of YLE Participants

Baseline Characteristics	N=1634	%
Cohort		
Cohort 1 – February 2023	109	6.7
Cohort 2 – March 2023	171	10.5
Cohort 3 – April 2023	167	10.2
Cohort 4 – May 2023	157	9.6
Cohort 5 – June 2023	119	7.3
Cohort 6 – July 2023	136	8.3
Cohort 7 – August 2023	191	11.7
Cohort 8 – September 2023	143	8.8
Cohort 9 – October 2023	152	9.3
Cohort 10 – November 2023	69	4.2
Cohort 11 – December 2023	220	13.5
Level of Education		
High School	77	4.7
Some College	145	8.9
Associate degree	143	8.8
Bachelor’s Degree	491	30
Master’s Degree	391	23.9
Doctorate Degree	68	4.2
Missing/Preferred not to respond	319	19.5
Race		
URM	267	16.3
Non-URM	1200	73.4
Missing/Preferred not to respond	167	10.2
Gender		
Man	419	25.6
Woman	1079	66
Non-Binary	6	0.4
Missing/Preferred not to respond	130	8
Mode		
In-person	755	46.2
Virtual	879	53.8

Table 4.6. Demographic Characteristics of YLE Sample Participants

Baseline Characteristics	N=160	%
Cohort		
Cohort 1 – February 2023	14	8.8
Cohort 2 – March 2023	18	11.3
Cohort 3 – April 2023	24	15.0
Cohort 4 – May 2023	16	10.0
Cohort 5 – June 2023	20	12.5
Cohort 6 – July 2023	8	5.0
Cohort 7 – August 2023	18	11.3
Cohort 8 – September 2023	12	7.5
Cohort 9 – October 2023	13	8.1
Cohort 10 – November 2023	11	6.9
Cohort 11 – December 2023	6	3.8
Level of Education		
High School	8	5.0
Some College	12	7.5
Associate’s Degree	12	7.5
Bachelor’s Degree	51	31.9
Master’s Degree	44	27.5
Doctorate Degree	6	3.8
Missing/Preferred not to respond	27	16.9
Race		
URM	27	16.9
Non-URM	121	75.6
Missing/Preferred not to respond	12	7.5
Gender		
Man	34	21.3
Woman	117	73.1
Non-Binary	0	0.0
Missing/Preferred not to respond	9	5.6
Mode		
In person	74	46.2
Virtual	86	53.8

LFC Program Evaluation Demographics

LFC Total Participants

Five LFC programs were held in 2023. Overall, 579 participants attended these programs. 55.3% of them participated in the in-person programs in March, June, and August, while 44.7% of them the virtual version in April and October. The average age of LFC participants was 43.5, ranging from 20 to 78. Most participants had a bachelor's or master's degree (62.9%) and identified as non-URM (76.9%), while participants from URM formed 14%. Most of the participants were female (67.2%). Like YLE, participants were at various career stages and worked in various sectors (i.e., governmental, private, non-profit, education, and faith). LFC participant demographics are presented in Table 4.7.

Table 4.7. Demographic Characteristics of LFC Participants

Baseline Characteristics	N=579	%
Cohort		
Cohort 1 – March 2023	60	10.4
Cohort 2 – April 2023	80	13.8
Cohort 3 – June 2023	123	21.2
Cohort 4 – August 2023	134	23.1
Cohort 5 – October 2023	182	31.4
Level of Education		
High School	31	5.4
Some College	48	8.3
Associate’s Degree	42	7.3
Bachelor’s Degree	187	32.3
Master’s Degree	177	30.6
Doctorate Degree	29	5
Missing/Preferred not to respond	65	11.2
Race		
URM	81	14
Non-URM	445	76.9
Missing/Preferred not to respond	53	9.1
Gender		
Man	144	24.9
Woman	389	67.2
Non-Binary	1	0.2
Missing/Preferred not to respond	45	7.8
Mode		
In-person	320	55.3
Virtual	259	44.7

LFC Sample Participants

90 participants responded to LFC T1 (before the program) and T2 (three months after) surveys. The March cohort had the largest response rate while the April cohort had the smallest. The average age for this pool was 46.4, with the youngest 24 and the eldest 75. Most participants had a bachelor’s or master’s degree (63.8%) and identified as non-URM (82.2%), while URM constituted 11.1%. Most of the participants were female (69.4%). There were a few participants

who preferred not to respond to demographic questions, particularly education level and race.

Gender was not asked consistently across all cohorts. Participant demographics can be found in Table 4.8.

Table 4.8. Demographic Characteristics of LFC Sample Participants

Baseline Characteristics	N=90	%
Cohort		
Cohort 1 – March 2023	18	20
Cohort 2 – April 2023	9	10
Cohort 3 – June 2023	15	16.7
Cohort 4 – August 2023	24	26.7
Cohort 5 – October 2023	24	26.7
Level of Education		
High School	3	3.3
Some College	9	10
Associate degree	7	7.8
Bachelor’s Degree	32	35.5
Master’s Degree	29	32.3
Doctorate Degree	4	4.4
Missing/Preferred not to respond	6	6.7
Race		
URM	10	11.1
Non-URM	74	82.2
Missing/Preferred not to respond	6	6.7
Mode		
In-person	57	63.3
Virtual	33	36.7

Focus Group Demographics

Two focus group sessions were conducted during the scale validation process. The first session held right after the content validity test, focused on discussing the content and face validity of the items. The second session, conducted towards the end, involved labeling the theoretical factors that emerged in EFA and confirmed during CFA. Participants for both groups were chosen from the panel of experts, with a specific emphasis on KLC’s teaching team (Table

4.9 and Table 4.10). Knowing that KLC is a key partner and main beneficiary of this research and will use the assessment tools for evaluating their future programs, priority was given to engaging KLC teachers in these focus groups.

Table 4.9. Focus Group 1 Participants

Experts	Background	Connection to Adaptive Leadership Framework
Participant 1	Theater, English, Leadership	Teacher-Coach-Practitioner
Participant 2	Organizational Management & Leadership	Teacher-Practitioner
Participant 3	Leadership	Teacher-Coach
Participant 4	Language & History (M.A)	Practitioner
Participant 5	Political Science	Teacher-Coach
Participant 6	Communication Studies (PhD)	Instructor-Teacher-Coach
Participant 7	Organizational Leadership	Teacher
Participant 8	Educational Leadership (PhD)	Instructor-Teacher-Coach

Table 4.10. Focus Group 2 Participants

Experts	Background	Connection to Adaptive Leadership Framework
Participant 1	Theater & English (BA)	Teacher-Coach-Practitioner
Participant 2	Language & History (M.A)	Practitioner
Participant 3	Political Science	Teacher-Coach
Participant 4	Communication Studies (PhD)	Instructor-Teacher-Coach
Participant 5	Organizational Leadership	Teacher
Participant 6	Psychology (PhD)	Teacher-Coach
Participant 7	Political Science	Practitioner

Interview Study Demographics

In response to our invitation to interview requested in the T4 survey, 10 participants expressed interest in being interviewed. In two rounds of interviews, these participants shared their successful and failed leadership practices during the four months after the YLE program. This pool of participants includes people from different walks of life and sectors. Table 4.11. Shows the demographics of the interviewees.

Table 4.11. Demographic Characteristics of the Interviewees

Baseline Characteristics		N=10	
Positions		Age	Pseudonyms Given
Academic Affairs Coordinator	1	-	Julia
Non-profit Director	1	39	Sydney
Community Theatre Producer	1	41	Ashley
Accessibility Board Member	1	49	Kait
Business Owner	1	74	Mary
Student Services Director	1	49	Tamara
Faith Network Team Leader	1	74	Bob
Health Coordinator	1	55	Kerry
CEO	1	58	Angie
Organizational Director	1	60	Tim
Level of Education			
Associate's Degree	1		
Bachelor's Degree	3		
Master's Degree	5		
Missing/Preferred not to respond	1		
Race			
URM	2		
Non-URM	8		
Gender			
Man	2		
Woman	8		

Demographics Highlights

There are several discernible aspects of the demographic characteristics of the sample participants, interviewees, and the focus groups' participants.

- Majority of survey participants were female and from non-URM groups.
- Participants come from different sectors (Private, education, business, faith, and non-profit)

- Participants with a variety of backgrounds from language and rhetoric to political science and psychology attended the focus groups. This was very important in accomplishing the focus group tasks mainly ensuring the face validity of items and labeling constructs.
- Interviewees were from different organizational positions: top-level managers, middle managers, team leaders, and volunteers.

Quantitative Findings by Hypotheses

Hypothesis 1: ALB Scale Validation

Developing new assessment tools for adaptive leadership in an engaged process was this study's first objective to address two main gaps in the leadership development process. The first gap, *teaching-assessment gap*, was addressed by creating assessment instruments tailored to adaptive leadership framework. The second gap, *measurement gap*, was tackled by establishing a collaboration between the researcher and KLC's LDP teaching and organizing team. Here, we describe how this objective was achieved.

Step 1: Literature Review & Extracting Scale Items

After reviewing Heifetz's books and articles about adaptive leadership — namely Heifetz (1994), Heifetz and Laurie (1997), Heifetz and Linsky (2002), Heifetz et al. (2009) — Heifetz et al. (2009) was selected to extract practical items describing adaptive leadership behavior. 47 items were extracted from this book (Appendix A.1) representing ten main domains: diagnose challenges, diagnose political landscape, manage self, energize others, intervene skillfully, act politically, engage courageously, run experiments, thrive, and orchestrate conflict. The scale items were written as sentences in a way that the authenticity of Heifetz's language be maintained. As such, most of the sentences were extracted directly with minor changes to make

the items independently readable and understandable. The researcher wanted to make items as straightforward as possible so that different interpretations could be avoided.

Step 2: Content and Face Validity Test

After extracting the items, a survey was conducted to test four CV criteria: essentiality (Likert scales 1-3), relevance, clarity, and simplicity (Likert scales 1-4). The survey was shared with a 21-member panel of experts who are teaching, researching, practicing, and coaching leadership through the lens of adaptive leadership (Table 4.1).

Per instructions in Chapter 3, the I-CV and S-CV scores as well as a probability of agreement chance were calculated for each scale item and the whole scale (Lynn, 1986; Polit & Beck, 2006). Out of 47 ALB scale items, 22 items passed the CV test, nine items were removed automatically due to low relevance (<0.7) (Lynn, 1986) or essentiality (conservative rate of <0.5), and 12 items were corrected based on survey suggestions. While two items needed minor revisions, 10 of these items had relevancy and essentiality values of above 0.78 and value above 0.9 for all the three simplicity, clarity, and ambiguity criteria (Appendix C). 14 items were sent to the focus group discussion for final decision and potential wording revision with high relevancy (>0.7) and medium essentiality (≥ 0.5 , <0.78) and low simplicity, clarity, and ambiguity. Although items ALB 18 and 25 had low essentiality, they were sent for final decision due to high relevancy. Nine items were sent to the focus group only for wording revision. These items had relevancy and essentiality values of above 0.7 with a value below 0.9 for at least one of the simplicity, clarity, and ambiguity criteria. Some exceptions might need special attention in focus group discussions to remove availability bias triggers (e.g., items 4, 32, 38). The researcher took a very conservative approach in removing the items solely based on the survey results.

An online 60-minute focus group discussion was designed after the CV test (see the session design in Appendix J). Eight members of the panel of experts who are KLC's teachers participated in the discussion to make final decision about the list of items and improving face validity. The researcher facilitated the discussion through the process discussed in Chapter 3. At the end of this section, 12 items were consensually removed by the panel. Moreover, several scale items were edited to become clearer and simpler and not to trigger the participants' availability biases with the KLC concepts. For instance, words, like diagnose, experiment, adaptive challenge, heat, my part of the mess, and comfort zone, were replaced with *analyze*, *test*, *complex challenge*, *intensity*, "*my contribution to challenge*", and "*the way people usually think and act*", respectively. Finally, the participants reached an agreement on the face validity of the items and a final list of 30 items (Appendix B.1).

Step 3: Exploratory Factor Analysis (EFA)

After the CV test and the focus group discussion, the agreed list of items underwent an EFA. After removing missing data and outliers ($-3 > Z$, $Z > +3$), EFA was conducted on 249 survey responses to investigate the underlying factor structure of the collected data. The first round of EFA was conducted on 150 data at the beginning of the study. However, per Kyriazos' (2018) recommendation, the researcher decided to increase the sample size to above 200. Expanding the dataset a few months after the first EFA did not change the factor structure and added to the reliability of EFA results. The EFA aimed to identify latent factors that explain the variance in the ALB. Sample adequacy, sphericity, factor loading, fit indices, and internal reliability are discussed in the following sections respectively (see Appendix Table D.6 for descriptives).

Sampling Adequacy. KMO measure of sampling adequacy was computed to assess the suitability of the data for factor analysis. The KMO values for all items are presented in Appendix Table D.6. Except items ALB24 and ALB30, all other items have KMO values above 0.7 showing that the data were suitable for EFA. Item ALB24 is kept due to its theoretical importance.

Sphericity. Bartlett's Test of Sphericity is statistically significant ($\chi^2 [435] = 2470, p < 0.001$), which supports the factorability of the data.

Extraction and Rotation Method. Principal Axis was used for initial factor extraction. To determine the number of factors, parallel analysis and the scree plot were used. To improve the interpretability of the factors, shorten the list of items, and especially due to the lack of confirmed theoretical constructs, an oblique rotation method of Promax was applied. We used this method based on the assumption that different adaptive leadership behaviors are correlated (Osborne, 2019). The rotated factor loadings, along with the eigenvalues, were examined to identify the strength and pattern of associations between the factors and the observed variables. Initially, six factors arose out of which the first five factors were kept for Scree Plot analysis to simplify the complexity of the model based on both elbow/breakpoint identification and theoretical explanation. Factor six was removed as its first and last items (ALB 30 & ALB12) had negative factor loadings and the rest two items (ALB28 & ALB 29) were not theoretically consistent.

The selected five factors together accounted for 36.65% of the total variance in the ALB data (Appendix Table D.2). Following scree plot analysis and theoretical considerations, factors 2,3, and 4 were combined into one factor and finally three factors were proposed for the CFA round of analysis (Appendix Figure D.1).

Factor Loading. Appendix Table D.1 presents the factor loadings for each item after rotation. Only loadings exceeding an absolute value of 0.35 are considered significant and interpreted. Item ALB1 & ALB27 were conservatively removed due to considerable cross-loadings on factors 1 and 2 (-0.357 & 0.395). Moreover, item ALB23 was kept and ALB16 was removed to ensure the theoretical consistency of the factors. Table 4.12 presents the final list of items after EFA.

Overall Fit Model. The chi-square test resulted in a p-value of less than 0.001 indicating good model fit. Other model fit tests like TLI and RMSEA were also in a satisfactory range (TLI=0.889 & RMSEA=0.0450) (see Appendix D).

Table 4.12. Component Tables and the ALB Items Kept

Factors	Items	
1	ALB8	I identify stakeholders and gather information about what they are afraid of giving up.
	ALB7	I identify stakeholders and gather information about their power or influence.
	ALB10	I consider multiple perspectives about problems at work/within the community.
	ALB23	I try new things to test my ideas.
	ALB15	I try to learn about different stakeholders related to the issue in my organization/community.
	ALB11	I engage others in considering multiple perspectives about problems at work/ within the community.
	ALB17	I implement my intervention ideas with small steps.
	ALB26	I regularly run tests to make progress on difficult challenges.
2	ALB3	I engage others in conversations about our organization's/community's conflicting priorities.
	ALB2	I discuss my conflicting priorities with others.
	ALB5	I engage others to say what is usually unspeakable in the organization/community.
	ALB4	I engage myself in conversations in which my competing commitments are discussed.
	ALB6	I call attention when my organization/community resists confronting challenges.
3	ALB13	I consult with others before taking action.
	ALB14	I let people in the organization/community work with my idea and give me feedback.
	ALB24	I allow for silence in meetings, arguments, and discussions.

Step 4: Confirmatory Factor Analysis (CFA)

CFA was conducted on two different datasets. Below the CFA results for each group of data are presented.

Group 1 CFA (337 Responses). 337 people responded to this round of survey after removing outliers ($Z < -3$, $Z > +3$) and responses with more than two missing items. The CFA was conducted to test the hypothesized factor structure identified in EFA and to assess the proposed model's fit. The analysis aimed to confirm whether the items load onto the specified latent factors per the EFA. The sections below summarize the key findings of the CFA. Moreover, see Appendix Table E.5 for descriptives.

Model Specification. First, the hypothesized model based on the EFA results was tested to see if it could be confirmed. While the model fit indices meet the minimum standards (see Appendix E), the researcher decided to test another potential model as an extension of this model as suggested by Lim and Jahng (2019). Following their recommendation to test a model with $K+1$ factors, a 3+1 factor model was tested as a candidate for estimating a number of factors. To do so, Factor 1 was subdivided into two factors: one group of items exhibited loadings above 0.6, while the other group displayed loadings below 0.5, indicating a substantial gap between them both in loading strength and theoretical conceptualization. Moreover, ALB 10 was removed due to its significant impact on the fit indices in both models.

Model Fit Indices (Model 2). The p-value of the chi-square ($\chi^2 [84] = [201]$, $p < 0.001$) indicated a good model fit. The RMSEA resulted in a value of 0.0644, suggesting an acceptable fit. The CFI value of 0.931 and TLI value of 0.914 showed excellent fit. The SRMR was found to be 0.0486 suggesting an excellent fit. The factor covariances did not show a high correlation between the factors. However, Structural Equation Modeling (SEM) was run to explore any

available higher-order model. SEM did not result in a new model with a better model fit.

Therefore, the four-factor structure model was adopted as the final model.

Factor Loadings. Appendix Table F.1 presents the standardized factor loadings for each item in the Model 2. Significant factor loadings ($p < 0.001$) indicate the strength and direction of the relationships between the latent factors and the items. All items have standard factor loading above 0.5 except ALB24 which was kept for its important contribution to the theoretical justification. This item describes an important behavior of a self-transforming mind. Items ALB7 and ALB15 have the highest factor loading. In this dataset, no high inter-factor correlation was observed (Appendix Table F.2).

Internal Consistency. Cronbach's α of 0.858 showed high internal consistency and reliability of the items in Model 2. The reliability statistics show that only dropping ALB24 will slightly increase α (Appendix Table F5 & F6). This item was not removed due to its importance to the theoretical justification of the model and the strength of the factors.

Group 2 CFA (581 Responses). The CFA was conducted for another dataset (second half of the year) to test the hypothesized factor structure identified in EFA and cross-validate the results of the first group to increase the generalizability. The sections below summarize the key findings of the CFA for Dataset 2. Moreover, see Appendix Table H.7 for Dataset 2 descriptives.

Model Specification. The approach on the first dataset was employed for the second dataset as well. Two models were run to do the CFA. Model 1 was conducted based on the EFA hypothesized three-factor structure (Appendix G). Similar to the first dataset, while the model fit indices were acceptable, the $K+1$ model resulted in significantly better fit indices. To the extent that it can be argued Model 2 is a very strong model.

Model Fit Indices (Model 2). The model was assessed based on the fit indices introduced in Chapter 3. The p-value of the chi-square (χ^2 [84] = 295, $p < .001$) indicates a good model fit. The RMSEA resulted in a value of 0.0657, suggesting an acceptable fit. The CFI value of 0.924 and TLI value of 0.905 show excellent fit. The SRMR was found to be 0.0481 suggesting excellent fit (Appendix H).

Factor Loadings (Dataset 2-Model 2). Appendix Table H1 presents the standardized factor loadings for each item in the CFA model for the second dataset. Significant factor loadings indicate the strength and direction of the relationships between the latent factors and the items ($p < .001$). Items ALB 13 and ALB24 were kept in the model due to their theoretical importance even though they have the least factor loading.

Internal Consistency. Cronbach's α of 0.86 shows high internal consistency and reliability of the items. The reliability statistics show that only dropping ALB24 and ALB13 will slightly increase α (Appendix Tables H5 and H6). These two items were not removed due to their importance to the theoretical justification of the model and the strength of the factors.

Model Selection. The initial specified model confirmed the three-factor structure suggested in EFA. However, after modifying the model following the Comparison Data approach suggested by Auerswald and Moshtagan (2019) and Lim and Jahng (2019), Factor 1 was split into two factors (Factor 1 and Factor 4) in both datasets. This improved model fit indices significantly enhanced the interpretability of factors, and more importantly, provided better theoretical justification.

Step 5: Construct & Criterion-Related Validity Tests

The construct validity of the newly developed ALB scale was examined to assess its ability to accurately measure the intended construct and to ensure that the scale measures what it

is expected and designed to measure. Criterion-related validity was also tested to assess the ability of the ALB scale to measure a target construct at current and future times. The sections below describe the results of the construct validity test for the ALB scale. For descriptives, see Appendix Table E.5 and H.7.

Convergent Validity. To test this validity, we used the same two datasets for CFA with 333 and 576 responses. The researcher examined the extent to which the ALB scores are positively correlated with Carless et al.'s (2000) TLB scale. The researcher asked ALB questions alongside the TLB questions in the surveys. The results confirmed this hypothesis, as a moderate positive correlation was observed between ALB and the TLB scales in both datasets (Ratner, 2009). The correlation coefficient in the two data sets is $r_1=0.587$ and $r_2=0.632$ respectively ($p < 0.001$) which indicates significant convergent validity (Appendix I).

Discriminant Validity. Discriminant validity was also assessed by examining the correlation between the ALB score and Diener's (1984) SJS as unrelated constructs in two datasets with 332 and 577 participants. This was conducted to ensure that the new scale does not yield spurious correlations with irrelevant constructs. The results in both datasets showed significant weak correlations between the ALB scale and SJS with $r_1=0.11$ ($p < 0.05$) and $r_2=0.196$ ($p < 0.001$), proving the discriminant validity of the new scale (Appendix I).

Concurrent Validity. It was hypothesized that scores on the ALB scale would positively correlate with scores on Hiller's (2005) LI at the current time. To test this hypothesis, we only used the second dataset with 575 responses as we did not ask concurrent validity questions in our first dataset. As a result, we observed a significant moderate positive correlation between the ALB scale and LI. The correlation coefficient is $r=0.413$ ($p < 0.001$) showing significant concurrent validity (Appendix I).

Predictive Validity. We conducted a regression analysis to examine the predictive validity of the ALB at T1 on LI at T3. A sample of 146 data points were used to run the regression analysis (See Appendix Table U.1.6 for descriptives). LI data was normalized using Box-Cox transformation as we found the skewness a little high (-0.985). The analysis controlled for individual differences by including relevant covariates such as sector, education level, URM/non-URM, and gender. Results revealed a significant positive relationship between ALB scores at T1 and LI at T3 ($\beta = 10.5$, $p < 0.001$ / Intercept= -16.6 , $p < 0.05$), indicating that higher levels of ALB at T1 were associated with higher LI at T3. This association remained robust even after accounting for individual differences and demonstrated a good fit (Adjusted R-squared = 0.21). These findings provide empirical support for the predictive validity of the ALB Scale in forecasting LI outcomes, underscoring its potential utility as a valuable tool for organizations seeking to enhance leader identity (Appendix U.1).

Step 6: Labeling Constructs: Interpreting Factors

After the CFA, items, and factors were taken to a second focus group discussion to be interpreted and labeled collectively by the panel of experts. Again, the KLC teaching team was invited to this focus group and engaged in the labeling process as described in Chapter 3 and outlined in Appendix J. Seven participants attended this 90-minute focus group on Zoom (KLC's director of research, chief civic leadership development officer, program manager, director of programs, senior director of learning and development, director of communications, and director of custom teaching and coaching).

Participants discussed the labels one by one and shared critiques and suggestions for each label. For some factors, participants decided to select one among the suggested ones while in others the groups co-created a new one. Table 4.13. shows the final labels after group discussion

and reaching a consensus. All participants agreed that the factors represent important concepts within the adaptive leadership framework.

Factor 1. Primarily consists of variables ALB7, 8, 11, and 15. The focus group labeled this factor as “*seek stakeholders’ perspectives*”.

Factor 2. Primarily consists of variables ALB2, 3, 4, 5, and 6. The focus group labeled this factor as “*Surface Conflict*”.

Factor 3. Primarily consists of items ALB 13, 14, and 24. The focus group labeled this factor as “*make space for multiple perspectives*”.

Factor 4. Primarily consists of variables ALB 17, 23, and 26. The focus group labeled this factor as “*Act experimentally*”.

Table 4.13. ALB Factors Labels

Factors	Leadership Behavior Items	Literature Domain	Labels
1	ALB8 I identify stakeholders and gather information about what they are afraid of giving up.	Diagnosing political landscape	Group 1: Diagnose across factions.
	ALB7 I identify stakeholders and gather information about their power or influence.	Diagnosing political landscape	Group 2: Identify stakeholders’ perspectives.
	ALB15 I try to learn about different stakeholders related to the issue in my organization/community.	Design effective interventions	Engaging stakeholders
	ALB11 I engage others in considering multiple perspectives about problems at work/ within the community.	Engaging others	Selected: Seek stakeholders’ perspectives.
2	ALB3 I engage others in conversations about our organization's/community's conflicting priorities.	Diagnosing challenges	Group 1: Make Conflicting Interpretations
	ALB2 I discuss my conflicting priorities with others.	Diagnosing challenges	Making space for others
	ALB5 I engage others to say what is usually unspeakable in the organization/community.	Diagnosing challenges	Group 2: Name competing values.
	ALB4 I engage myself in conversations in which my competing commitments are discussed.	Diagnosing challenges	
	ALB6 I call attention when my organization/community resists confronting challenges.	Diagnosing challenges	Selected: Surface conflict

3	ALB13	I consult with others before taking action.	Design effective interventions	Group 1: Create/Make Space multiple perspectives Group 2: Collaboration Seeking others' influence Selected: Make space for multiple perspectives.
	ALB14	I let people in the organization/community work with my idea and give me feedback.	Design effective interventions	
	ALB24	I allow for silence in meetings, arguments, and discussions.	Deploy yourself	
4	ALB26	I regularly run tests to make progress on difficult challenges.	Run experiments	Group 1: Test Interpretation Group 2: Experimental mindset Selected: Act experimentally
	ALB23	I try new things to test my ideas.	Engage courageously	
	ALB17	I implement my intervention ideas with small steps.	Deploy yourself	

Hypothesis 1: ALSE Scale Validation

Step 1: Literature Review & Extracting Scale Items

After reviewing Heifetz's books and articles about adaptive leadership, namely Heifetz (1994), Heifetz and Laurie (1997), Heifetz and Linsky (2002), and Heifetz et al. (2009), 11 items were extracted directly (Appendix A.2) representing ten main behaviors: diagnose challenges, diagnose political landscape, manage self, energize others, intervene skillfully, acting politically, engage courageously, run experiments, thrive, and orchestrate conflict. The scale items were written as sentences in a way that the authenticity of Heifetz's language be maintained. To keep this authenticity, the researcher focused on the main competencies of adaptive leadership and avoided developing a long list of items. As such, most of the sentences were extracted directly with minor changes to make the items independently readable and understandable. The researcher wanted to make items as straightforward as possible so that different interpretations could be avoided.

Step 2: Content and Face Validity Test

After extracting the items, a survey was conducted to test four content validity (CV) criteria: essentiality (Likert scales 1-3), relevance, clarity, and simplicity (Likert scales 1-4). The survey was shared with a 21-member panel of experts who are teaching, researching, practicing, and coaching leadership through the lens of adaptive leadership (Table 4.1)

The I-CVI and S-CVI scores as well as a probability of agreement chance were calculated per instructions in Chapter 3 for each scale item and the whole scale (Lynn, 1986; Polit & Beck, 2006). Out of 11 ALB scale items, 10 items passed the CV test, one item was removed automatically due to low relevance and essentiality (<0.78) (Lynn, 1986), and five items were corrected based on the suggestions on the survey. These items had I-CVIs above 0.9 in all of the five criteria. Two items were sent to the same focus group discussion held for ALB only for wording revision (items ALSE1 & 3) (Appendix C). These two items had minimum relevancy and essentiality of 0.78 and value below 0.9 for at least one of the simplicity, clarity, and ambiguity criteria. The researcher took a very conservative approach in removing the items.

Step 3: Pilot Study: Exploratory Factor Analysis (EFA)

After the content validity test and the focus group discussion, the agreed list of items underwent a pilot study to conduct the EFA. After removing missing data and the outliers ($-3 > Z$, $Z > +3$), EFA was conducted on 241 survey responses to investigate the underlying factor structure of the collected data. This analysis aimed to identify latent factors that explain the variance in the ALSE. Sample adequacy, sphericity, factor loading, fit indices, and internal reliability are discussed in the following sections respectively. Moreover, see Appendix Table D.6 for descriptives.

Sampling Adequacy. KMO measure of sampling adequacy was computed to assess the suitability of the data for factor analysis. The KMO value for all items can be found in Appendix K. All items have a KMO value above 0.8 showing that the data were suitable for EFA.

Sphericity. Bartlett's Test of Sphericity is statistically significant ($\chi^2 [45] = 1436$, $p < 0.001$), which supports the factorability of the data (Appendix K.).

Extraction and Rotation Method. The Principal Axis was used for initial factor extraction as the Shapiro-Wilk normality test showed a little deviation from normality. To determine the number of factors, parallel analysis and the scree plot (Appendix Figure K.1) were used. Two factors were kept, which accounted for 61.6% of the total variance in the ALSE data. To improve the interpretability of the factors and especially because of the lack of confirmed theoretical constructs, an oblique rotation method of Oblimin was applied. We used this method based on the assumption that different adaptive leadership competencies are correlated (Osborne, 2019). The rotated factor loadings, along with the eigenvalues, were examined to identify the strength and pattern of associations between the factors and the observed variables. Using “Promax Rotation” did not change the factor structure.

Factor Loading. Appendix Table K.1 presents the factor loadings for each item after rotation. Only loadings exceeding an absolute value of 0.35 are considered significant and interpreted. ALSE8 and ALSE6 with high cross-loading values of 0.4 were removed. The final list accepted for the CFA round can be found in Table 4.14.

Overall Fit Model. The chi-square test resulted in a p-value of less than 0.001 indicating good model fit. Other model fit tests like TLI and RMSEA were also in a satisfactory range (TLI=0.912 & RMSEA=0.131). This model is proposed for the CFA round of analysis.

Table 4.14. ALSE Components and the Items Kept

Factors	Items	
1	ALSE3	I can analyze the root causes of a challenge.
	ALSE1	I can analyze situations containing complex challenges.
	ALSE2	I can analyze how my organization/community is responding to difficult challenges.
	ALSE10	I can test ideas to make progress on challenges.
2	ALSE4	I can manage myself amidst uncertainty and conflict.
	ALSE5	I can energize others to make progress on difficult challenges.
	ALSE9	I feel that I know how to inspire others.
	ALSE7	I am confident of my ability to manage conflict.

Step 4: Main Study: Confirmatory Factor Analysis (CFA)

CFA was conducted on two different datasets. Below the CFA results for each group of data are presented.

Group 1 CFA (333 Responses). After removing missing data and the outliers from the dataset, the CFA was conducted on 333 responses to test the hypothesized factor structure identified in EFA and to assess the proposed model's fit. The analysis aimed to verify whether the items load onto the specified latent factors in accordance with the EFA results. The sections below summarize the key findings of the CFA. Moreover, see Appendix Table E.5 for descriptives.

Model Specification. The CFA model was specified with two latent factors and their corresponding items. The model was tested against the null hypothesis that the proposed factor structure is not a good fit for the data. Removing items ALSE 4 and 10 significantly improved the model fit indices making the model very strong.

Model Fit Indices. The model was assessed based on the fit indices introduced in Chapter 3. The p-value of the chi-square ($\chi^2 [8] = [21.7]$, $p < 0.01$) indicates a good model fit. The RMSEA resulted in a value of 0.0717, suggesting an acceptable fit. The CFI value of 0.985

and TLI value of 0.971 show excellent fit. The SRMR was found to be 0.0340 suggesting an excellent fit.

Factor Loadings. Appendix Table L.1. presents the standardized factor loadings for each item in the CFA model. Significant factor loadings indicate the strength and direction of the relationships between the latent factors and the items ($p < 0.001$). All items have factor loading above 0.6 with ALSE5 (0.852) the highest standard estimate and ALSE7 the lowest (0.648). Factor estimates do not show a high correlation between the two factors (Appendix Table L.2).

Internal Consistency. Cronbach's α of 0.854 shows high internal consistency and reliability of the items. The reliability statistics show that droppings will not increase α .

Group 2 CFA (548 Responses). The CFA was conducted for another set of data to test the hypothesized factor structure identified in EFA and cross validate the results of the first group to increase the generalizability of the results (Appendix M). The sections below summarize the key findings of the CFA for Dataset 2. Moreover, see Appendix Table H.7 for Dataset 2 descriptives.

Model Specification. The CFA model again was specified with two latent factors and their corresponding items. The model was tested against the null hypothesis that the proposed factor structure is not a good fit for the data. Like dataset 1, removing ALSE 4 and 10 significantly improved the model fit.

Model Fit Indices. The model was assessed based on the fit indices introduced in Chapter 3. The p-value of the chi-square ($\chi^2 [8] = [30.3]$, $p < 0.001$) indicates a good model fit. The RMSEA resulted in a value of 0.0712, suggesting an acceptable fit. The CFI value of 0.986 and TLI value of 0.973 show excellent fit. The SRMR was found to be 0.0307 suggesting an excellent fit.

Factor Loadings. Appendix Table M.1 presents the standardized factor loadings for each item in the CFA model for the second set of data. Significant factor loadings indicate the strength and direction of the relationships between the latent factors and the items ($p < 0.001$). The minimum factor loading belongs to ALSE7 (0.689) while ALSE3 in this dataset has the highest standard estimate.

Internal Consistency. Cronbach's α of 0.872 shows high internal consistency and reliability of the items. Appendix Table M.6 shows data dropping will not increase the internal consistency rate.

Model Selection. The model presented here is a modified version of what is suggested in EFA. The model demonstrated a much better model fit (particularly regarding RMSEA and TLI) after removing items ALSE 4 and 10. Appendix Table M.4 shows how the model fit indices change after removing these three items. However, we suggest keeping item ALSE10 as it improves the theoretical validity of the scale and removing it will not affect the model fit considerably especially in Dataset 2.

Step 5: Construct & Criterion-Related Validity Tests

The construct validity of the newly developed ALSE scale was examined to assess its ability to accurately measure the intended construct and to ensure that the scale measures what it is expected and designed to measure. Criterion-related validity was also tested to assess the ability of the ALSE scale to measure a target construct at current and future times. The sections below describe the results of the construct validity test for the ALSE scale. For descriptives, see Appendix Table E.5 and H.7.

Convergent Validity. To test this validity, again, we used two sets of data with 327 and 542 responses. The researcher examined the extent to which the ALSE scores are positively

correlated with Murphy's (1992) MSE. The researcher asked ALSE questions alongside the MSE questions in the surveys. The results confirmed this hypothesis, as a strong positive correlation was observed between ALSE and Murphy's (1992) SE scales in both data sets. The correlation coefficient in the two datasets is $r_1=0.796$ and $r_2=0.777$ respectively ($p < 0.001$) which indicates significant convergent validity (Appendix N).

Discriminant Validity. Discriminant validity was also assessed by examining the correlation between the ALSE and SJS as unrelated constructs. This was conducted in two datasets with 325 and 537 participants to ensure that the new scale does not yield spurious correlations with irrelevant constructs. The results showed significant positive low correlations between the ALSE scale and SJS with $r_1=0.251$ ($p < 0.001$) and $r_2=0.186$ ($p < 0.001$) respectively, proving the discriminant validity of the new scale (Appendix N).

Concurrent Validity. It was hypothesized that scores on the ALSE scale would positively correlate with scores on TLB at the current time. To test this hypothesis, we again used two data sets with 325 and 537 participants. As a result, we observed a significant moderate positive correlation between the ALSE scale and TLB. The correlation coefficient in the two datasets is $r_1=0.505$ and $r_2=0.495$ respectively ($p < 0.001$) which indicates significant concurrent validity (Appendix N).

Predictive Validity. We conducted a regression analysis to examine the predictive validity of the ALSE at T1 on ALB at T3. A sample of 150 data points were used to run the regression analysis (See Appendix U.2.6 for descriptives). The analysis controlled for individual differences by including relevant covariates such as sector, education level, URM/non-URM, and gender. Results revealed a significant positive relationship between ALSE scores at T1 and ALB

at T3. However, the very small ($\beta = 0.01$, $p < 0.001$) indicates that higher levels of ALSE at T1 might not strongly be associated with higher levels of ALB (Appendix U.2).

Step 6: Labeling Constructs: Interpreting Factors

The ALSE factors were labeled in the same focus group as ALB's. Table 4.15. shows the preliminary labels and final labels for Factors 1 and 2 selected by the focus group.

Factor 1. Primarily consists of variables ALSE1, 2, 3, and 10. The focus group labeled this factor as “*Confidence to diagnose*”.

Factor 2. Primarily consists of variables ALSE5, ALSE7, ALSE9. The focus group labeled this factor as “*Confidence to energize others*”.

Table 4.15. ALSE Factors Labels

Factors	Self-Efficacy Items		Labels
1	ALSE3	I can analyze the root causes of a challenge.	Group 1: I can fully understand the situation. Confidence to diagnose. Group 2: Individual capacity to analyze and intervene. Selected: Confidence to diagnose.
	ALSE1	I can analyze situations containing complex challenges.	
	ALSE2	I can analyze how my organization/community is responding to difficult challenges.	
	ALSE10	I can test ideas to make progress on challenges.	
2	ALSE5	I can energize others to make progress on difficult challenges.	Group 1: Confidence to act Group 2: Confidence in energizing others Selected: Confidence to energize others.
	ALSE9	I feel that I know how to inspire others.	
	ALSE7	I am confident of my ability to manage conflict.	

Hypothesis 1: ALSA Scale Validation

Step 1: Literature Review & Extracting Scale Items

After reviewing Heifetz's books and articles about adaptive leadership-namely Heifetz (1994), Heifetz and Laurie (1997), Heifetz and Linsky (2002), Heifetz et al. (2009)- Heifetz et

al.'s (2009) book was selected to extract practical items describing ALSA. 25 items were extracted directly from this book (Appendix A.3). The scale items were written as sentences in a way that the authenticity of Heifetz's language be maintained. To keep this authenticity, the researcher focused on the main capacities of adaptive leadership and avoided developing a long list of items. As such, most of the sentences were extracted directly with minor changes to make the items independently readable and understandable. The researcher wanted to make items as straightforward as possible so that different interpretations could be avoided.

Step 2: Content and Face Validity Test

After extracting the items, a survey was conducted to test four content validity (CV) criteria: essentiality (Likert scales 1-3), relevance, clarity, and simplicity (Likert scales 1-4). The survey was shared with a 21-member panel of experts who are teaching, researching, practicing, and coaching leadership through the lens of adaptive leadership (Table 4.1). As Appendix C shows, the number of experts who responded to all survey questions varies. There are two major reasons behind this. First, due to the long length of the survey, some experts did not respond to all items' CV questions. Second, since the essentiality question was added a few days after the first survey was submitted, not all experts responded to this question. However, the panel's size met the standard for all items' CV questions, though (Lynn, 1986; Yusoff, 2019).

The I-CV and S-CV scores as well as a probability of agreement chance were calculated per instructions in Chapter 3 for each scale item and the whole scale (Lynn, 1986; Polit & Beck, 2006). Out of 25 ALSA scale items, 19 items passed the CV test, six items were removed automatically due to low relevance (<0.78) (Lynn, 1986) or essentiality (conservative rate of <0.5), and eight items were corrected based on survey suggestions. Six of these items had I-CVIs above 0.9 in all of the five criteria and two of them needed minor revisions. Nine items were sent

to the focus group discussion for final decision and potential wording revision with high relevancy (>0.78) and medium essentiality (≥ 0.5 , <0.78) and low simplicity, clarity, and ambiguity, and two items were sent to focus group only for wording revision. These two items had minimum relevancy and essentiality of 0.78 and a value below 0.9 for at least one of the simplicity, clarity, and ambiguity criteria (ALSA3, 4, 11) (Appendix Table C.3). The researcher took a very conservative approach in removing the items.

Step 3: Pilot Study: Exploratory Factor Analysis (EFA)

After the content validity test and the focus group discussion, the agreed list of items underwent a pilot study to run the EFA. After removing missing data and outliers, EFA was conducted on 248 survey responses to investigate the underlying factor structure of the collected data. This analysis aimed to identify latent factors that explain the variance in the ALSA. Sample adequacy, sphericity, factor loading, fit indices, and internal reliability are presented in Appendix O and discussed in the following sections respectively. Moreover, see Appendix Table D.6 for descriptives.

Sampling Adequacy. KMO measure of sampling adequacy was computed to assess the suitability of the data for factor analysis. The KMO value for all items can be found in Appendix Table O.5. All items have a KMO value above 0.6 showing that the data were suitable for EFA.

Sphericity. Bartlett's Test of Sphericity is statistically significant ($\chi^2 [91] = 598$, $p < 0.001$), which supports the factorability of the data. This means that the data can be used to identify underlying factors.

Extraction and Rotation Method. Maximum Likelihood was used for initial factor extraction due to the normality of the data. Changing the extraction method to “Principal Axis” or “Minimum Residuals” did not change the factor structure. To determine the number of factors,

parallel analysis and the scree plot (Appendix Figure O.1) were used. Three factors were kept, which accounted for 31% of the total variance in the ALSA data (Appendix Table O.2) To improve the interpretability of the factors and especially because of the lack of confirmed theoretical constructs, an oblique rotation method of Oblimin was applied. We used this method based on the assumption that different adaptive leadership behaviors are correlated (Osborne, 2019). The rotated factor loadings, along with the eigenvalues, were examined to identify the strength and pattern of associations between the factors and the observed variables. It is worth noting that, using other extraction and rotation methods including orthogonal ones did not change the factor structure.

Factor Loading. Appendix Table O.1 presents the factor loadings for each item after rotation. Only loadings exceeding an absolute value of 0.35 are considered significant and interpreted. Items ALSA12,13, and 4 with loading below 0.35 were removed from the list.

Overall Fit Model. The chi-square test resulted in a p-value of less than 0.05 indicating good model fit. Other model fit tests like TLI and RMSEA were also in a satisfactory range (TLI=0.924 & RMSEA=0.0409).

Table 4.16. ALSA Component and the Items Kept

Factors		Items
1	ALSA6	6. I am aware of my level of tolerance during moments of chaos, conflict, and confusion in my organization/community.
	ALSA8	8. I am aware of the way I usually think and act.
	ALSA7	7. I am aware of the triggers that make me angry/overwhelmed.
	ALSA9	9. I am aware of the distinction between my role (e.g., colleague, partner, friend, sibling) and who I am.
2	ALSA1	1. I know what role I play in relation to an important challenge in my organization/community.
	ALSA11	11. I have strong reason(s) to practice leadership.
	ALSA2	2. I regularly assess my motivations for trying to create change.
	ALSA10	10. I am aware of what I am formally authorized to do within my organization/community.
3	ALSA3	I know which of my behaviors I need to change to lead more effectively.
	ALSA5	I am aware of my unmet personal needs that could impede my exercise of leadership.

Step 4: Main Study: Confirmatory Factor Analysis (CFA)

CFA was conducted on two different sets of data. The CFA results for each group of data are presented in Appendices P and Q and discussed below.

Group 1 CFA (347 Responses). 347 people responded to this survey after removing outliers and missing data. The CFA was conducted on these two datasets to test the hypothesized factor structure identified in EFA and to assess the proposed model's fit. The analysis aimed to verify whether the items load onto the specified latent factors in accordance with the EFA. The sections below summarize the key findings of the CFA. Moreover, see Appendix Table E.5 for descriptives.

Model Specification. The CFA model was specified with two latent factors and their corresponding items. The model was tested against the null hypothesis that the proposed factor structure is not a good fit for the data.

Model Fit Indices. The model was assessed based on the fit indices introduced in Chapter 3. For the first dataset, the p-value of the chi-square ($\chi^2 [32] = 85.8, p < <0.001$) indicates good model fit. The RMSEA resulted in a value of 0.0696, suggesting an acceptable fit. The CFI value of 0.915 and TLI value of 0.880 show excellent fit. The SRMR was found to be 0.0519 suggesting an excellent fit.

Factor Loadings. Appendix Table P.1. presents the standardized factor loadings for each item in the CFA model. Significant factor loadings indicate the strength and direction of the relationships between the latent factors and the items ($p < 0.001$). ALSA5 has the lowest factor loading (0.43) while ALSA7 has the highest (0.7). Factor covariances do not show any high correlation between the factors (Appendix Table P.2).

Internal Consistency. Cronbach's α of 0.755 shows high internal consistency and reliability of the items. The reliability statistics show that dropping items will not increase α .

Group 2 CFA (602 Responses). The CFA was conducted for another set of data to test the hypothesized factor structure identified in EFA and cross-validate the results of the first group to increase the generalizability of the results. The sections below summarize the key findings of the CFA for Dataset 2. Moreover, see Appendix Table H.7 for Dataset 2 descriptives.

Model Specification. The CFA model again was specified with two latent factors and their corresponding items. The model was tested against the null hypothesis that the proposed factor structure is not a good fit for the data.

Model Fit Indices. The model was assessed based on the fit indices introduced in Chapter 3. The p-value of the chi-square (χ^2 [32] = 94.2, $p < .001$) indicates a good model fit. The RMSEA resulted in a value of 0.0568, suggesting an acceptable fit. The CFI value of 0.937 and TLI value of 0.912 show excellent fit. The SRMR was found to be 0.0406 suggesting an excellent fit.

Factor Loadings. Appendix Table Q.1. presents the standardized factor loadings for each item in the CFA model for the second set of data. Significant factor loadings indicate the strength and direction of the relationships between the latent factors and the items ($p < .001$). All items have factor loading above 0.5 except ALSA11 (0.46) and ALSA10 (0.48). No correlation was observed between the factors (Appendix Table Q.2).

Internal Consistency. Cronbach's α of 0.752 shows high internal consistency and reliability of the items. The reliability statistics show that dropping will not increase the α .

Model Modification. No model modification is needed compared to EFA.

Step 5: Construct & Criterion-Related Validity Tests

The construct validity of the newly developed ALSA scale was examined to assess its ability to accurately measure the intended construct and to ensure that the scale measures what it is expected and designed to measure. Criterion-related validity was also tested to assess the ability of the ALSA scale to measure a target construct at current and future times. The sections below describe the results of the construct validity test for the ALSA scale. For descriptives, see Appendix Table H.7.

Convergent Validity. To test this validity, again, we used one dataset with 345 responses. The researcher examined the extent to which the ALSA scores are positively correlated with Neider and Schreisheim's (2011) NSA as the sister scale. The researcher asked ALSA questions alongside the SA questions in the survey. The results confirmed this hypothesis, as a moderate positive correlation was observed between ALSA and the SA scales in both data sets. The correlation coefficient is $r=0.552$ ($p < 0.001$) which indicates significant convergent validity.

Discriminant Validity. Discriminant validity was also assessed by examining the correlation between the ALSA and Allen and Meyer's (1990) PCS as unrelated constructs. This was conducted in a dataset with 586 participants to ensure that the new scale does not yield spurious correlations with irrelevant constructs. The results showed a significant positive low correlation between the ALSA scale and the PCS with $r=0.101$ ($p < 0.05$), proving the discriminant validity of the new scale.

Concurrent Validity. It was hypothesized that scores on ALSA scale would positively correlate with scores on TLB at current time. To test this hypothesis, we again used two datasets with 332 and 572 participants. As a result, we observed a significant moderate positive

correlation between the ALSA scale and TLB in both datasets. The correlation coefficient in the two data sets is $r_1=0.555$ and $r_2=0.547$ respectively ($p < 0.001$) which indicates significant concurrent validity.

Predictive Validity. We conducted a regression analysis to examine the predictive validity of the ALSA at T1 on ALB at T3. A sample of 150 data points was used to run the regression analysis (See Appendix Table U.2.6 for descriptives). The analysis controlled for individual differences by including relevant covariates such as sector, education level, URM/non-URM, and gender. Results revealed a significant positive relationship between ALSA at T1 and ALB at T3 ($\beta = 0.26$, $p < 0.001$), indicating that higher levels of ALSA at T1 were associated with higher ALB at T3. This association remained robust even after accounting for individual differences and demonstrated a good fit (Adjusted R-squared = 0.076). These findings provide empirical support for the predictive validity of the ALSA Scale in forecasting ALB outcomes, underscoring its potential utility as a valuable tool for organizations seeking to enhance adaptive leadership behavior (Appendix U.3).

Step 6: Labeling Constructs: Interpreting Factors

The ALSA factors were labeled in the same focus group as ALB and ALSE except for Factor 3. Factor 3 emerged as a strong factor later when the researcher added more data to the EFA dataset. Hence, Factor 3 was not discussed in the focus group and was labeled by the researcher himself. Table 4.17. shows the preliminary labels and final labels selected by the focus group for Factor 1 and Factor 2.

Factor 1. Primarily consists of variables ALSA6, 7, 8, and 9. The focus group labeled this factor as “*Awareness of defaults*”.

Factor 2. Primarily consists of variables ALSA1, 2, 10, and 11. The focus group labeled this factor as “*Awareness of role & purpose*”.

Factor 3. Primarily consists of variables ALSA3 and 5. The researcher labeled this factor as “*Awareness of needs*”.

Table 4.17. ALSA Factors Labels

Factors		Self-Awareness Items	Literature Domains	Labels
1	ALSA6	I am aware of my level of tolerance during moments of chaos, conflict, and confusion in my organization/community.	Personal tuning	Group 1: Awareness of defaults
	ALSA8	I am aware of the way I usually think and act.	Bandwidth	Group 2: Self-awareness
	ALSA7	I am aware of the triggers that make me angry/overwhelmed.	Personal tuning	
	ALSA9	I am aware of the distinction between my role (e.g., colleague, partner, friend, sibling) and who I am.	My roles	Selected: Awareness of defaults
2	ALSA1	I know what role I play in relation to an important challenge in my organization/community.	Me as a system	Group 1: Purpose and role clarity
	ALSA11	I have strong reason(s) to practice leadership.	Purpose	Group 2: Awareness of role
	ALSA2	I regularly assess my motivations for trying to create change.	Me as a system	
	ALSA10	I am aware of what I am formally authorized to do within my organization/community.	My roles	Selected: Awareness of purpose and role
3	ALSA3	I know which of my behaviors I need to change to lead more effectively.	Me as a system	Selected: Awareness of needs
	ALSA5	I am aware of my unmet personal needs that could impede my exercise of leadership.	Personal tuning	

Hypothesis 1: ALSR Scale Validation

Step 1: Literature Review & Extracting Scale Items

After reviewing Heifetz's books and articles about adaptive leadership and Stanovich and West's (2011) taxonomy of thinking errors and listed thinking dispositions. Seven items were developed describing the reflective mind's potential commands for exercising leadership consisting of action orientations and purposes (Appendix A.4). Developing ALSR in this research largely served as an initial attempt to determine if such a component distinct from MRM- is involved in decision making for adaptive intervention. Due to time limits and restrictions of survey length, the ALSR validation process did not necessarily follow the best practices' validation process, especially the item generation phase. The purpose here is to make the case for developing a stronger ALSR in the future.

Step 2: Content and Face Validity Test

After developing the items, a survey was conducted to test four content validity (CV) criteria: essentiality (Likert scales 1-3), relevance, clarity, and simplicity (Likert scales 1-4). The survey was shared with a 21-member panel of experts who are teaching, practicing, and coaching leadership through the lens of adaptive leadership (Table 4.1)

The I-CV and S-CV scores were calculated per instructions in Chapter 3 for each scale item and the whole scale. All seven ALSR items passed the CV test. Four items were corrected based on the suggestions on the survey and none of them needed to be discussed in the focus group discussion, Hence, the final list was derived from the content validity test and sent to the EFA study (Appendix B.4).

Step 3: Pilot Study: Exploratory Factor Analysis (EFA)

After removing missing data and outliers, EFA was conducted on 245 survey responses to investigate the underlying factor structure of the collected data. This analysis aimed to identify latent factors that explain the variance in the ALSR. Sample adequacy, sphericity, factor loading, fit indices, and internal reliability are discussed in the following sections respectively. Moreover, see Appendix Table D.6 for descriptives.

Sampling Adequacy. KMO measure of sampling adequacy was computed to assess the suitability of the data for factor analysis. The KMO value for all items is presented in Appendix Table R. 5. All items have a KMO value above 0.55 showing that the data met the minimum suitability for EFA.

Sphericity. The Bartlett's Test of Sphericity is statistically significant ($\chi^2 [21] = 171$, $p < <0.001$), which supports the factorability of the data.

Extraction and Rotation Method. Maximum Likelihood was used for initial factor extraction due to normality of the data. To determine the number of factors, parallel analysis and the scree plot were used. Two factors were kept, which accounted for 29% of the total variance in the ALSR data (Appendix Table R.2.). To improve the interpretability of the factors, and due to lack of confirmed theoretical constructs as well as uncertainty about the correlation of factors, the oblique rotation method of Promax was applied. The rotated factor loadings, along with the eigenvalues, were examined to identify the strength and pattern of associations between the factors and the observed variables. It is worth noting that, using other extraction and rotation methods including orthogonal ones did not change the factor structure.

Factor Loading. Appendix Table R.1. presents the factor loadings for each item after rotation. Since the number of items is low, we kept all items for the CFA analysis even though

items ALSR7 and 6 have loadings below 0.35. Moreover, due to ALSR7 cross-loading in both factors, two models will be tested in CFA to see in what factor SR7 fits better. Table 4.18 presents the final items kept for CFA.

Overall Fit Model. The chi-square test resulted in a p-value of less than 0.001 indicating good model fit. Other model fit tests like TLI and RMSEA were also in a satisfactory range (TLI=0.868 & RMSEA=0.0615).

Table 4.18. Component Tables and the ALSR Items Kept

Factors	Items	
1	ALSR1	Creating change at work/in my community is my personal priority.
	ALSR2	Creating change is part of my job description.
	ALSR3	I practice leadership because a person should always consider new ways of solving a problem.
	ALSR6	Even if my work/community environment had been different, I probably would have had the same tendency for change.
	ALSR7	I believe there is little I can do to change the way things are.
2	ALSR4	I believe we should look to people in positions of authority to solve problems.
	ALSR5	Difficulties can usually be overcome through waiting for good fortune, rather than thinking about the problem.
	ALSR7	I believe there is little I can do to change the way things are.

Step 4: Main Study: Confirmatory Factor Analysis

CFA was conducted on a dataset with 601 survey responses. The CFA results are presented in Appendix S and discussed below.

Model Specification. The CFA model again was specified with two latent factors and their corresponding items showing that SR7 fits better in Factor 1. The model was tested against the null hypothesis that the proposed factor structure is not a good fit for the data.

Model Fit Indices. The model was assessed based on the fit indices introduced in Chapter 3. The p-value of the chi-square ($\chi^2 [8] = 31, p < 0.001$) indicates a good model fit. The RMSEA resulted in a value of 0.0692, suggesting an acceptable fit. The CFI value of 0.94 and

TLI value of 0.887 show excellent fit. The SRMR was found to be 0.0416 suggesting an excellent fit.

Factor Loadings. Appendix Table S.1. presents the standardized factor loadings for each item in the CFA model for the second dataset. Significant factor loadings indicate the strength and direction of the relationships between the latent factors and the items ($p < 0.001$). ALSR1 has the highest factor loading while ALSR7 has the lowest. As expected, no strong correlation was observed between the two factors (Appendix Table S.2).

Internal Consistency. Cronbach's α of 0.574 shows moderate internal consistency and reliability of the items. The reliability statistics show that dropping items ALSR1, ALSR4, and ALSR7 will slightly increase the α . These items were not removed due to their importance to the theoretical justification of the model and the strength of the factors.

Model Modification. Removing ALSR6 improved the model fit indices significantly. Moreover, while ALSR7 fits better in Factor 1, we are not still sure about what exact factor it belongs to. The generalizability test (Appendix T) showed that in some populations (e.g., government and non-profit sectors) this scale presents a better model fit having ALSR7 under the second factor. Moreover, theoretically, this item aligns more with Factor 2.

Step 5: Construct & Criterion-Related Validity Tests

The construct validity of the newly developed ALSR scale was examined to assess its ability to accurately measure the intended construct and to ensure that the scale measures what it is expected and designed to measure. Criterion-related validity was also tested to assess the ability of the ALSR scale to measure a target construct at current and future times. The sections below describe the results of the construct validity test for the ALSR scale. For descriptives, see Appendix Table H.7.

Convergent Validity. To test this validity, we used one dataset with 591 responses. The researcher examined the extent to which the ALSR scores are positively correlated with ALSA validated in this research. This was more encouraged after the qualitative interviews and the phenomenological analysis that showed goal orientation dispositions and self-awareness are important drivers of adaptive leadership activities. The researcher asked ALSR questions alongside the ALSA questions in the surveys. The results confirmed this hypothesis, as a moderate positive correlation was observed between ALSA and the ALSR scales in both data sets (Hinkel et al., 2003). The correlation coefficient is $r=0.407$ ($p < 0.001$) which indicates acceptable convergent validity (Appendix N).

Discriminant Validity. Discriminant validity was also assessed by examining the correlation between the ALSR and MRM as unrelated constructs. It was assumed that these two types of thinking dispositions do not have a bidirectional relationship. The qualitative analysis also showed that people with a tendency to make multiple simulations may not necessarily tend to intervene and vice versa. This was conducted in a dataset with 597 participants to ensure that the new scale does not yield spurious correlations with irrelevant constructs. The results showed a significant positive low correlation between the ALSR scale and the MRM with $r=0.344$ ($p < 0.001$), proving the discriminant validity of the new scale (Appendix N).

Concurrent Validity. It was hypothesized that scores on the ALSR scale would positively correlate with scores on ALB at the current time. To test this hypothesis, we used a dataset with 578 participants. As a result, we observed a significant moderate positive correlation between the ALSR scale and ALB. The correlation coefficient is $r=0.532$ ($p < 0.001$) which indicates significant concurrent validity (Appendix N).

Predictive Validity. We conducted a regression analysis to examine the predictive validity of the ALSR at T1 on ALB at T3. A sample of 150 data points was used to run the regression analysis (See Appendix Table U.2.6 for descriptives). The analysis controlled for individual differences by including relevant covariates such as sector, education level, URM/non-URM, and gender. Results revealed a significant positive relationship between ALSR at T1 and ALB at T3 ($\beta = 0.27$, $p < 0.001$), indicating that higher levels of ALSR at T1 were associated with higher ALB at T3. This association remained robust even after accounting for individual differences and demonstrated a good fit (Adjusted R-squared = 0.094). These findings provide empirical support for the predictive validity of the ALSR Scale in forecasting ALB outcomes, underscoring its potential utility as a valuable tool for organizations seeking to enhance adaptive leadership behavior (Appendix U.4).

Step 6: Labeling Constructs: Interpreting Factors

The ALSR factors were labeled in a different small focus group asynchronously via Google Doc. This happened due to time limits during the second focus group. Four participants from the second focus group participated in this asynchronous labeling. Table 4.19 shows the small group's preliminary labels and final labels selected.

Factor 1. Primarily consists of variables ALSR1, 2, 3, 7. The selected label is *“Purposefulness”*.

Factor 2. Primarily consists of variables ALSR4 and 5. The selected label is *“Dependence on authority”*.

Table 4.19. ALSA Factors Labels

Factors	Items	Labels
1	ALSR1 Creating change at work/in my community is my personal priority.	1. Self-authorization, agency
	ALSR2 Creating change is part of my job description.	2. Purposefulness
	ALSR3 I practice leadership because a person should always consider new ways of solving a problem.	Selected: Purposefulness
	ALSR7 I believe there is little I can do to change the way things are.	
2	ALSR4 I believe we should look to people in positions of authority to solve problems.	1. Dependence on authority
	ALSR5 Difficulties can usually be overcome through waiting for good fortune, rather than thinking about the problem.	2. Lack of purposefulness and agency -Authority dependence
		Selected: Authority Dependence

Hypothesis 1: New Scales' Generalizability Test

To test the generalizability of the factor structure, several different rounds of CFA were conducted on different groups of participants based on time (Before the program and 30 days after) as well as social criteria including race (URM and non-URM), gender (Man, Woman – Non-binary data was not sufficient to run the analysis), age groups (Gen X, Millennials, and Gen Z), sector (education, business, government/politics, nonprofit) (Appendix T.1). The datasets used in the previous section was broken into these subgroups and CFA was run again to see if the model shows consistency. Data shows that the factor structure has been consistent for all measurements across almost all of these groups. There were a few inconsistencies that are worth exploring in the future research. Appendix T list the inconsistencies with specific examples. Below, a few of these inconsistencies are highlighted,

- In some groups removing particular items improved the model. For instance, in Women's category, removing ALSR's second factor improved the model fit significantly. Also, for government sector, removing ALSA 3 & 10 improved the model.

- Gen Z data showed several inconsistencies. For instance, ALSE model changed to a one-factor model for this group. Also, ALB13 (*I consult with others before taking action*) was negatively correlated with the scale. While this result is interesting and merits further investigation, additional data, especially from different populations and contexts, can help determine whether this result is consistent or merely noise.
- ALSA factor structure significantly changed for Gen Z and business sector categories.
- ALSR did not converge in some categories such as master's degree and Gen Z group.

While at this stage we might not be able to explain why these particular inconsistencies occurred, one general hypothesis is that some items/factors might seem more/not discerning/important in particular groups. Future studies might help us with calibrating these measurement tools for different groups and contexts. Moreover, the generalizability test is not limited to only these groups, and it is suggested that the measurement tools be tested in other aspects of Cronbach et al.'s (1953) generalization theory such as cultural contexts other than Kansas.

Hypothesis 3: Relationship Between Leadership Competencies and Reflective Mind

Linear Mixed Model (LMM)

Hypothesis 3 was examined here before Hypothesis 2 as the analytical approach used in this hypothesis sets the foundation for addressing the second hypothesis.

Regression models were run separately at two timepoints (T1, T3) to investigate the relationship between ALB and reflective mind components (ALSE, ALSA, MRM, and ALSR) explored in this research based on Stanovich's tripartite model. Although the regression models at T1 and T3 depicted a statistically significant relationship between the predictor variables — ALSE, ALSR, ALSA, MRM — and ALB, the residuals' correlation suggests that the model is

not efficient and practical. This correlation implies that the regression model might not be capturing the variability of the relationship between predictor and outcome variables and there is room for more complex models. We did not want to ignore this correlation as it might have significant impact on the estimators of the regression model (Martin, 1992; Thoresen & Laake, 2007). Hence, per Haslett and Hayes's (1998) recommendation about considering marginal and conditional residuals, we decided to conduct a model that captures the effects that are not explained in the regression model: linear mixed model (LMM).

As discussed in Chapter 3, LMMs are used when data are collected according to repeated measures design. In such designs, there would likely be correlations in variables since surveys are designed to be asked of the same participants at different timepoints. Using Jamovi software and Gallucci's (2019) GAMLj package, LMM was used to increase the accuracy of the model by correct modeling of longitudinal correlation and including random effects.

After removing missing data and outliers, 159 responses from T1-T3 and 84 Responses matched across T1, T3, and T4 were selected to run the LMMs. Among these responses, there were a few missing data that were presented in Appendix Table V.2.9. These missing data were not removed from the matched responses as LMM handles these data through maximum likelihood estimation (Rosenkranz, 2015).

Before running the model, each predictor variable was centered by subtracting the grand mean from individual observations. The centering procedure ensures that the coefficients represent the change in the dependent variable associated with a one-unit change in the centered predictors while holding other variables constant. The models were created using the Maximum Likelihood method due to its advantages in providing unbiased fixed effects estimates and the

comparison essence of our analysis using LRT besides other fit indices per Meteyard and Davies (2020) instructions.

All components were added meticulously while monitoring the changes in fit indices and coefficient estimates and their associated standard errors and p-values. Fixed effects of the reflective mind components were added incrementally based on this finding in the qualitative phase that the first step in the decision-making process is to deal with cognitive dissonance. Hence, ALSA was first included in the model as a self-regulatory process as Day et al. (2009) would describe, or as an emotional regulatory process according to Stanovich et al.'s (2011) list of fluid rationality components. ALSA captures a person's awareness of defaults, role and purpose, and needs which helps with holding to purpose in the anxiety-filled moment after dissonance and even after disrupting the status quo which is the following step.

The second step would be stopping the automatic system of mind initiated by the first cognitive decoupling command of the reflective mind. Then, the mind shifts to the second decoupling process which is making multiple simulations. These processes are initiated by a reflective mind component that here was measured via MRM. Anchored in our theoretical framework, MRM includes cognitive decoupling 1 which overrides the status quo, and cognitive decoupling 2 which updates the status quo model with a new representational model. Moving to act on the simulations needs a disposition to continue simulations (cognitive decoupling 3). This was measured by ALSR and ALSE.

In LMM the order of adding IVs does not affect the estimate and the explanatory power of the model as we used maximum likelihood estimation. The order described above in adding the components into the LMM model only provides the logic behind numbering the models. From a theoretical perspective, although we describe these processes and their associated

measurement tools here in a linear way, the decision-making process to exercise leadership might involve these elements in parallel or in an interconnected way. For instance, in the qualitative findings, we will show that “awareness of defaults” might come before MRM while “awareness of purpose and role” might come after it. Showing the exact order of the reflective mind process falls outside the scope of this research.

Four LMM scenarios with 10 models were specified progressively to determine how the variability of the dependent variable (ALB) can be attributed to fixed and random effects: Null, basic, intermediate, and advanced scenarios. We started with a simple model including random effects first (Null scenario). Then, reflective mind components as predictor variables were added to the model separately (Basic scenario). Afterward, the predictor variables were added incrementally and in interaction terms (Intermediate scenario). Finally, random slopes were included in the model to reach the highest level of complexity (Advanced Scenario).

Null Scenario (Model 1). To start, we calculated the average ALB across all participants. Yet, participants enter the program residing at different levels of leadership behavior. To consider these differences and each participant’s baseline of leadership behavior, we modeled the random variance in intercepts (Appendix V.1).

The null model reveals that the random effect of participants’ intercept on the ALB is statistically significant (Estimate=5.22, $p < 0.001$, R-Squared Conditional=0.59). The model fitting process converged successfully, and the parameter estimates were obtained using the "bobyqa" optimizer as the default choice. Kolmogorov-Smirnov test, residual histogram, Q-Q, and random coefficient plots show the normality of residuals. Using different DF methods Satterthwaite and Kenward-Roger did not majorly change the results.

This model's output shows that the average variance associated with random intercepts is 0.27 and the grand mean through which participants' intercepts vary around is $SD=0.52$. This variation shows to what extent participants may differ in the level of their leadership behavior. In this case, we can consider SD as small. ICC is 0.59 which suggests that 59% of variability in ALB can be attributed to individual differences. The significance of random effect LRT ($p < 0.001$) suggests that a more complex model could be available. Hence, in the next section, we decided to conduct a basic scenario by adding the fixed effects to the model one by one (See Appendix V).

Basic Scenario (Models 2-5). The fixed effects of the predictor variables (ALSR, MRM, ALSE, ALSA) were isolated in the model to capture their explanatory power. All the models converged successfully. ALSA explained the biggest variation in ALB (18%) and MRM the least with about 9%. The coefficients were estimated with associated standard errors and p-values. Kolmogorov-Smirnov test, residual histogram, Q-Q, and random coefficient plots show the normality of residuals which indicates the efficiency of parameter estimates. Using different DF methods Satterthwaite and Kenward-Roger did not majorly change the results. The significance of LRT indicates that there is room for more complex models (See Table 4.20).

Intermediate Scenario (Models 6-9). In this scenario, the fixed effects of the predictor variables (ALSR, MRM, ALSE, ALSA) were added in interaction terms and incrementally in an order based on Stanovich's tripartite model of reasoning process (ALSA, MRM, ALSE & ALSR). Adding each predictor variable to the null model incrementally improved the explanatory strength of the model, reduced the distance between $R^2(C)$ and $R^2(M)$, and resulted in significant estimates (Table 4.21). Also, for each model, the LRT significance suggested more complexity can be tested. Moreover, among all potential interactions between the predictor

variables only the interaction between MRM and ALSE resulted in significant results (Estimate=0.008, $p<0.05$). LMM9 as the final model in this scenario consists of random intercepts, the fixed effects of all predictor variables, and the interaction between MRM and ALSE (Table 4.20). For more details see Appendix V.2.

The model fitting process converged successfully, and the parameter estimates were obtained using the "bobyqa" optimizer. The coefficients were estimated with associated standard errors and p-values. Kolmogorov-Smirnov test, residual histogram, Q-Q, and random coefficient plots show the normality of residuals which indicates the efficiency of parameter estimates. Using different DF methods Satterthwaite and Kenward-Roger did not majorly change the results. The final intermediate model shows that the fixed effect of all predictor variables is statistically significant, indicating a positive contribution to the variations in the dependent variable. The omnibus test shows that the inclusion of the fixed effects collectively contributes to explaining ALB significantly. The variables' p-values show that there is a statistically significant relationship between the IVs and ALB. ALSA has the largest coefficient estimate while ALSE and its interaction with MRM have the least. Removing ALSE from the model increases ALSR, MRM, and ALSA's estimate by 0.03, 0.05, and 0.08 units which confirms their moderate correlation calculated in the scale validation section. Removing ALSE reduces the model fit significantly. Therefore, it will be kept in the model. Random effects are also statistically significant with a variance of 0.168. ICC 0.580 indicates that around 58% of the total variance in the ALB can be attributed to the differences between individuals.

LMM8 and LMM9 are the most similar model to the regression model run separately for each timepoint. While the regression models did not converge, the efficiency of models 6-8 shows how adding individuals' random effects improves the practicality of the model. However,

the significance of LRT suggests that a more complex model can predict more variability of the dependent variable.

Advanced Scenario (Model 10). In this scenario, we added random slopes to LMM9. Only the interaction term's random slope was loaded on the model. The MRM*ALSE random slope indicates variability in its effect across individuals and tells us people may vary in how the interaction between MRM and ALSE influences their ALB levels (Appendix V.2).

The addition of this random slope resulted in an improvement in both the R^2 -Marginal and Conditional values and reduced the gap between them (Table 4.21). Furthermore, the non-significant LRT suggests that LMM10 does not significantly enhance the explanatory power compared to LMM9. However, since the inclusion of the MRM*ALSE slope leads to a slight improvement in the conditional R-squared value and the significance of interactions, a reduction in AIC, and the creation of the most complex, efficient, and theoretically aligned model, it is chosen as the final model. Additionally, the correlation between the random intercept and random slope offers insights into the nuanced relationship between MRM, ALSE, and ALB, which will be discussed later in this chapter. This model predicts 35.5% of the variance in ALB through its fixed effects and 73.6% through its fixed and random effects (Table 4.20).

Table 4.20. LMM Scenario Results

Model Fit (Converged models)	Null Scenario	Basic Scenario				Intermediate Scenario				Advanced Scenario
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
	ALB ~ 1 + 1 ID)	ALB ~ 1 + AL SA +(1 ID)	ALB ~ 1 + MRM +(1 ID)	ALB ~ 1 + ALSE +(1 ID)	ALB ~ 1 + ALSR +(1 ID)	ALB ~ 1 + ALSA + MRM + (1 ID)	ALB ~ 1 + ALSA + MRM + ALSE + (1 ID)	ALB ~ 1 + ALSA + MRM + ALSE +ALSR+(1 ID)	ALB ~ 1 + ALSA + MRM + ALSE +ALSR+MRM*ALS E+ (1 ID)	ALB ~ 1 + ALSA + MRM + ALSE +ALSR+MRM*ALSE+ (1 + MRM*ALSE ID)
AIC	686.337	591.862	651.5918	617.256	620.056	567.281	549.466	527.722	525.328	524.987
BIC	698.205	607.665	667.4048	633.08	635.879	587.021	573.154	555.358	556.912	564.468
LogLikel.	-340.169	-291.931	-321.7959	-304.628	-306.028	-278.641	-268.733	-256.861	-254.664	-252.494
R²- Marg.	0	0.182	0.0862	0.18	0.173	0.234	0.287	0.341	0.351	0.355
R²- Cond.	0.589	0.682	0.6247	0.653	0.598	0.705	0.726	0.730	0.728	0.736
ICC	0.589	0.611	0.589	0.577	0.514	0.615	0.616	0.591	0.580	0.597
LRT	105	107	106	102	69.1	110	113	95.1	90.5	4.34
LRT <i>p</i>-value	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.114

Table 4.21. LMMs Explanatory Strengths Change

Models	Effects Added Incrementally	R ² (M)	R ² (C)	ΔR ² (M)	ΔR ² (C)	R ² (C)-R ² (M)
1	Intercept + (1 ID)	0	0.589	—	—	0.589
2	Intercept + ALSA + (1 ID)	0.182	0.682	0.182	0.093	0.500
6	Intercept + ALSA + MRM + (1 ID)	0.234	0.705	0.052	0.023	0.471
7	Intercept + ALSA + MRM + ALSE + (1 ID)	0.287	0.726	0.053	0.021	0.439
8	Intercept + ALSA + MRM + ALSE + ALSR + (1 ID)	0.341	0.730	0.054	0.004	0.389
9	Intercept + ALSA + MRM + ALSE + ALSR + MRM*ALSE + (1 ID)	0.351	0.728	0.010	0.045	0.424
10	Intercept + MRM + ALSA + ALSE + ALSR + MRM*ALSE + (1+ MRM*ALSE ID)	0.355	0.736	0.004	0.008	0.381

Hypothesis 2: KLC's Path 2023 Impact Assessment

The YLE and LFC programs underwent a longitudinal evaluation study, which aimed to address the *training transfer gap* by examining changes in leadership competencies over time. As discussed in Chapter 1, assessing competencies immediately after the program may not provide a realistic view of its actual impact, as individuals often encounter difficulties in transferring training to practice. Longitudinal assessment has the potential to offer a more realistic view of the impact of LDPs when individuals return to work. Additionally, this evaluation study provided an opportunity to monitor the effectiveness of the newly developed assessment tools in practice.

YLE Impact

After matching participant data across four timepoints, removing the missing data, removing responses with unacceptable time duration (below six minutes), and deleting the outliers ($-3 < Z$, $Z > +3$), a sample of 159 participants (for T1-T3) and 84 participants for T4 was selected to measure the impact of YLE program on ALB. For measuring the variation in leadership behavior T1, T3, and T4 datasets were used as we did not ask ALB questions at T2. The LMM model was run for each leadership competency separately including the fixed effect of Time and random effects of individual variability. Time was scaled as a difference between each subsequent timepoint to capture the temporal progression and examine the changes over time. We did not restrict the Likelihood Model so that LRT shows us meaningful changes in the variability of fixed and random effects per Luke's (2017) guidance. The same process was applied for other competencies (MRM, ALSE, ALSR, and ALSA) at three timepoints (T1, T2, and T3) with 141, 137, 141, and 141 participants respectively (See Appendix W.6 for descriptives).

LFC Impact

After matching participant data across two timepoints (T1, before the program; and T2 (Three months after), removing the missing data, removing responses with unacceptable time duration (below six minutes), and deleting the outliers ($-3 < Z$, $Z > +3$), a sample of 90 participants was selected to measure the impact of LFC program on ALB and the reflective mind components. The LMM model was run for each variable separately including the fixed effect of Time and random effects of individual variability. Time was scaled as a difference between each subsequent timepoint to capture the temporal progression and examine the changes over time. We did not restrict the Likelihood Model so that LRT shows us whether the model is complex enough to capture the variability of fixed and random effects.

As participants attended LFC either individually or in their organizational groups, another LMM was also run to measure the nested group effect on the outcome and independent variables. We also were curious to see if the program has the same effect for each group and if the relationship between Time (attributed to the LFC program) and the ALB is the same across groups).

Result 1: ALB Increased Over Time After YLE Program

LMM exhibited a significant main effect of Time (R^2 -Marginal=0.045, $F=24.1$, $p<0.001$) showing changes in ALB across time even after controlling for baseline. Post-hoc comparison tests were conducted to explore the specific changes over time. Kolmogorov-Smirnov test, Residual histogram, Q-Q, and random coefficient plots show the normality of residuals. Results revealed a significant increase in ALB from T1 to T3 (Estimate = 0.28, $p < 0.001$) and a sustained effect at T4 (Estimate = 0.30, $p < 0.001$) (Appendix W.1).

We also conducted LRT to assess the overall significance of the Time's fixed effects in the model compared to the null model. The LRT results indicate that the inclusion of the fixed effects for Time significantly improves the model fit (LRT=129, $p < 0.001$). Thus, if we attribute the effect of Time here to the YLE program, we conclude that there is a statistically significant overall effect of the LDP on ALB.

Polynomial terms up to the second degree (quadratic) were also added to the model to capture the nonlinear trends. The inclusion of polynomial terms revealed a linear effect of T1-T3 on ALB (Estimate=0.22, $p < 0.001$) and a nonlinear negative effect of T1-T4 (Estimate=-0.1, $p < 0.01$). This provides a more nuanced understanding of the temporal dynamics in ALB which may distinguish a more initial impact of the program from the long-term effect, saturation in practicing leadership, or participants dividing into two major groups of experimenters and non-experimenters in the long term. The increase between T3 and T4 is not statistically significant which might be due to not enough data at this time ($p_{\text{bonferroni}} = 1$). (Appendix W.1).

Result 2: ALSA Increased Over Time After YLE Program

LMM exhibited a significant main effect of Time ($R^2\text{-Marginal}=0.1$, $F=54.1$, $p < 0.001$) showing changes in ALSA across time even after controlling for baseline. Post-hoc comparison tests were conducted to explore the specific changes over time. Kolmogorov-Smirnov test, Residual histogram, Q-Q, and random coefficient plots show the normality of residuals. Results revealed a significant increase in ALSA from T1 to T2 (Estimate = 0.42, $p < 0.001$). This increase was sustained even one month after the program (Estimate = 0.15, $p < 0.001$). There is a 0.06 unit decrease in the estimate between T2-T3 which is not statistically significant ($p_{\text{bonferroni}} = 1$).

We also conducted LRT to assess the overall significance of the Time's fixed effects in the model compared to the null model. The LRT results indicate that the inclusion of the fixed effects for Time significantly improves the model fit (LRT=119, $p<0.001$). Thus, if we attribute the effect of Time here to the YLE program, we conclude that there is a statistically significant overall effect of the LDP on ALSA.

Polynomial terms up to the second degree (quadratic) were also added to the model to capture the nonlinear trends. The inclusion of polynomial terms revealed a linear effect of T1-T2 on ALSA (Estimate=0.25, $p<0.001$) and a nonlinear negative effect of T1-T3 (Estimate=-0.2, $p<0.001$). This provides a more nuanced understanding of the temporal dynamics in ALSA which may distinguish a more initial impact of the program from the long-term effect or be a sign of saturation in awareness (Appendix W.2).

Result 3: MRM Increased Over Time After YLE Program

LMM exhibited a significant main effect of Time (R^2 -Marginal=0.008, $F=6.25$, $p<0.01$) showing changes in MRM across time even after controlling for baseline. Yet, this effect is too small. Post-hoc comparison tests were conducted to explore the specific changes over time. Kolmogorov-Smirnov test, Residual histogram, Q-Q, and random coefficient plots show the normality of residuals. Results revealed a significant increase in MRM from T1 to T2 (Estimate = 0.12, $p < 0.001$). This increase sustained even till T3 (Estimate = 0.08, $p=0.053$) (Appendix W.3).

Result 4: ALSR Increased Over Time After YLE Program

LMM exhibited a significant main effect of Time (R^2 -Marginal=0.03, $F=14.7$, $p<0.001$) showing changes in ALSR across time. Post-hoc comparison tests were conducted to explore the

specific changes over time. Results revealed a significant increase in ALSR from T1 to T2 (Estimate = 0.3, $p < 0.001$) and T1 to T3 (Estimate = 0.2, $p < 0.05$) (Appendix W.4).

Result 5: No Evidence about the Impact on ALSE

LMM exhibited a significant main effect of Time (R^2 -Marginal=0.016, $F=10.1$, $p<0.001$) showing changes in ALSE across time. Yet, this model is not efficient since there is a correlation between residuals. Hence, the role of the program in predicting variation in ALSE is in doubt.

Result 6: LFC Participants Report More ALB

Result 6.1. LMM exhibited a significant main effect of Time (R^2 -Marginal=0.08, $F=35$, $p<0.001$) showing changes in ALB across time. Post-hoc comparison tests were conducted to explore the specific changes over time. Results revealed a significant increase in ALB from T1 to T2 (Estimate = 0.33, $p < 0.001$) (Appendix X.1).

We also conducted LRT to assess the overall significance of the Time's fixed effects in the model compared to the null model. The LRT results indicate that the inclusion of the fixed effects for Time significantly improves the model fit ($LRT=36.9$, $p<0.001$). Thus, if we attribute the effect of Time here to the YLE program, we conclude that there is a statistically significant overall effect of the LDP on ALB.

Change in other competencies was also examined. LMM model with Time as the IV did not show any statistically significant increase in MRM. In contrast, ALSR increased between T1-T4 significantly (Estimate=15, $p<0.05$) as well as ALSA (Estimate=0.35, $p<0.001$). Moreover, unlike YLE, results show LFC's impact on ALSE across time (Estimate=5.13, $p<0.001$) (Appendix X2-5).

Result 6.2. The LMM analysis revealed a significant increase in ALB between T1 and T2 across both Group I, consisting of individuals who attended the LFC program individually, and

Group G, comprising individuals who attended the program in their organizational groups. While both groups demonstrated an increase in ALB over time, Group I exhibited a slightly higher level of ALB compared to Group G. However, it's important to note that the difference in ALB between the two groups did not reach statistical significance (Estimate Time (T2-T1)¹*Group (G-I) = -0.11, p = 0.3). This lack of statistical significance may be attributed to the relatively small sample size of 90 participants. Nonetheless, the general LMM analysis did yield a significant result indicating a higher level of ALB in Group I compared to Group G. If we attribute Time to the program, we can assert that, although both groups received the same training program and their leadership behavior overall increased, the program's impact on groups of individuals' ALB is more than organizational groups (Appendix X.6).

Result 7: YLE's Model of Impact

In testing Hypothesis 2, we developed 10 LMMs to explore the connection between reflective mind components and ALB. The most optimized and complex model demonstrated that a combination of fixed and random effects accounts for approximately 73% of the variance in ALB. Here, we aim to investigate the influence of KLC's intervention, specifically the YLE program, on this model. Our goal is to illustrate how the YLE program alters the dynamics of the model and identify the mechanisms through which it enhances individuals' ALB.

Basic Time Scenario (Model T_{Basic Model}). Time was added to the model to see the effect of Time (Here YLE program) on the ALB. While Time individually predicts the increase in ALB, Time's estimate (as an added predictor to other IVs) was not significant suggesting that the program may not have a uniform effect on ALB across different levels of other variables (Appendix Y.1).

Intermediate Time Scenario (Model T_{Intermediate}). In this scenario, we included predictor variables paired interactions with time in the model incrementally. Among all more than 10 potential models, the best fit with significant results belongs to Model T_{Intermediate} containing time interactions with ALSR, ALSE, and ALSA. Time's effect on ALB is consistent across different levels of MRM (Appendix Y.2).

Removing the main effect of Time from model T_{Basic} while adding it in interaction terms with ALSA and ALSR improved the model fit especially R-squared marginal and conditional as well as the power of coefficients' estimates and power of omnibus fixed effect tests. Moreover, the interaction between Time and ALSR, despite significance, resulted in a negative estimate which probably suggests collinearity between the ALSR's main effect and the interaction. In deciding which one to choose, we found keeping the ALSR interaction with Time improves the model while keeping the omnibus test for all fixed effects significant. Also, we removed ALSA's main effect from the model as controlling for it in the model showed collinearity with its interaction effect and did not change model indices. Hence, to make the model more parsimonious, ALSR and ALSA's effects were defined merely through their interaction with time. We also replaced ALSE with its time interaction. The model fit indices especially, AIC and BIC, improved slightly compared to the model with only ALSE's main effect (Appendix Y.2).

Model T_{Intermediate} converged successfully within an acceptable fit range, and the parameter estimates were obtained using the "bobyqa" optimizer. The coefficients were estimated with associated standard errors and p-values. Kolmogorov-Smirnov test, Residual histogram, Q-Q, and random coefficient plots show the normality of residuals which indicates the efficiency of parameter estimates. Using different DF methods Satterthwaite and Kenward-Roger did not majorly change the results. The final intermediate model shows that the fixed

effect of all predictor variables is statistically significant, indicating a positive contribution to the variations in the dependent variable. The omnibus test shows that the inclusion of the fixed effects collectively contributes to explaining ALB significantly.

The variables' p-values show that there is a statistically significant relationship between all the IVs and ALB and ALSR, ALSE, and ALSA interactions with Time. These interactions improved the model R-squared (Explanatory strength) significantly. This suggests that Time's (Here program's) effect on ALB is not uniform across all levels of ALSE, ALSA, and ALSR. In other words, the effect of the program on ALB is higher at higher levels of these predictor variables. Random effects are also statistically significant with a variance of 0.175. ICC 0.602 indicates that around 60% of the total variance in the ALB can be attributed to the differences between individuals.

Advanced Scenario (Model T_{Advanced}). In this model, all predictor variables' random slopes were added to Model T_{Intermediate}. Only the random slope of MRM's interaction with ALSE was loaded on the model. Similar to LMM10, while it did not improve AIC and BIC and R-squared marginal, it improved the R-squared conditional slightly. Hence, the final model consists of MRM and its interaction's random slope, Time interaction with ALSA, ALSR, ALSE, and fixed and random intercept. Overall, this model predicts 75.4% of the variation in adaptive leadership behavior through its fixed and random effects. It also predicts 36% of this variation through its fixed effects (Appendix Y.3).

Qualitative Research Findings

In the sequential explanatory approach, the qualitative research was conducted to explain one major why in the quantitative phase with more concrete and in practice details: The relationship between reflective mind components and adaptive leadership behavior. It does so by describing the decision-making process people take to exercise leadership. This decision-making process not only provides insights about the dynamics between reflective mind components and leadership intervention but also offers information about the different factors of assessment tools. This phase of the research contributes to addressing the training transfer gap discussed in chapter 1. Here, we present this decision-making process.

Visual Representation of Data

Two types of representations were used to make sense of the qualitative data. First, a code book was used to list the codes extracted from the transcripts. Then, concept mapping was employed as a technique to conduct the second cycle of coding and depict the relationship between these codes.

Code Mapping

NVivo was employed to do the code mapping. According to Anfara (2008), code mapping “brings meaning, structure, and order to data” (as cited in Saldaña, 2021, p. 297). Codes, memos, and early categories were organized in NVivo’s coding table. This basic outlining format made it possible to map versus, value, and emotion codes and facilitated the transition from the first to the second coding cycle (Table 4.22). The researcher also used memo notes to document the process codes.

Table 4.22. Code Book

Participants' Codes	Ref	Participants' Codes	Ref
<i>Board Member (Failed Action)</i>		<i>Theatre Producer (Failed Action)</i>	
Reasoning process		Actions after failed actions	2
Emotions at moment of non-action	1	Evaluation of failure	8
Impediments of action	3	Impediments of success	2
Dispositions	1	Intervention reaction	5
Types of leadership behavior		Reasoning process (Failed)	
Modelling behavior	2	Cognitive dissonance	5
Taking temperature	2	Emotion during the action	1
<i>Board Member (Successful Action)</i>		Emotion now about failure	1
Evaluation	1	Emotions after action	4
Others' reactions	2	Emotions after dissonance	1
Reasoning process		Evaluation reasoning	2
Disposition	6	Sticking to habits-role-authority	4
Emotions after action	3	Types of leadership behavior (Failed)	
Emotions after dissonance	3	Asking questions	1
Simulation of self-failure	1	Making not enough interpretations	2
Stick to authority mindset	5	YLE Contribution	2
Type of leadership behavior		<i>Theatre Producer (Successful Action)</i>	
Raising the heat	4	Evaluation	2
<i>Academic Affairs Coordinator (Failed Action)</i>		Learning from experience as motivation	1
Alternative action	3	Reasoning Process	
Environment	8	Cognitive dissonance	3
Impediments	1	Emotions after	6
Egos	1	Emotions before action	3
Risk	1	Emotions during action	1
Tension	8	Emotions now about actions	1
Reasoning process		Motivations to act	
Dispositions	1	Dispositions	8
Emotions at the moment	2	Override the autonomous system	2
Type of leadership behavior		Thoughts after action	1
No action (because of authority)	2	Self-Awareness	1
<i>Student Services Director (Failed Action)</i>		Types of leadership behavior	
Alternative actions	6	Making multiple interpretations	13
Reasoning process		Hypothetical thinking	1
Cognitive dissonance	9	Managing heat (conflict)	2
Emotions after dissonance	6	Role of Person- Asking questions	8
Intervention	3	YLE Contribution	
Motivations	3	Tools for bolstering dispositions	1
Others' reactions	1	Tools for managing self	2
<i>Student Services Director (Successful Action)</i>		<i>Business Owner (Failed Actions)</i>	
Leadership is difficult	1	Reasoning Process	
Mastery goal orientation	5	Active mind	3
Reasoning process		Disposition	3
Cognitive dissonance	2	Evaluation after action	2
Dispositions	4	Manage emotions	2
Emotions after action	1	<i>Business Owner (Successful Action)</i>	
Emotions during action	2	Reasoning Process	
impediments	2	Active mind	6
Intervention (Alone)	3	Dispositions	2
Motivation to act		Emotion before action	5

Participants' Codes	Ref	Participants' Codes	Ref
Innate	3	Type of leadership behavior	
My husband	1	Creating space for others	3
Personality	2	Managing conflict	9
Race	1	Seek perspectives	2
Self-questioning	5	Non-profit Director (Failed Action)	
Self-talk after dissonance	4	Actions after failed actions	1
Self-awareness		Impediments	1
Me and my role	1	Reasoning process	
CEO (Failed Action)		Emotions after failed action	2
Reasoning process		Emotions before action	1
Cognitive dissonance		Emotions of non-action anymore	2
Emotions after dissonance		Tiredness as impediment	3
Intervention		Types of leadership behavior	
Motivations		Failed Experiment	1
Others' reactions		Non-profit Director (Success)	
Types of leadership behavior		Emotions after action	1
Start seeking perspectives		Evaluation	2
Taking care of self		Intervention reactions	5
		Reasoning process	
		Disposition	2
		Emotion before action	1
		Emotion during action	1
		Types of leadership behavior	
		Creating space for perspectives	1
		Designing experiment	1
		Managing conflict	1
		Managing frustration	2
		Raise the heat (Conflict)	2

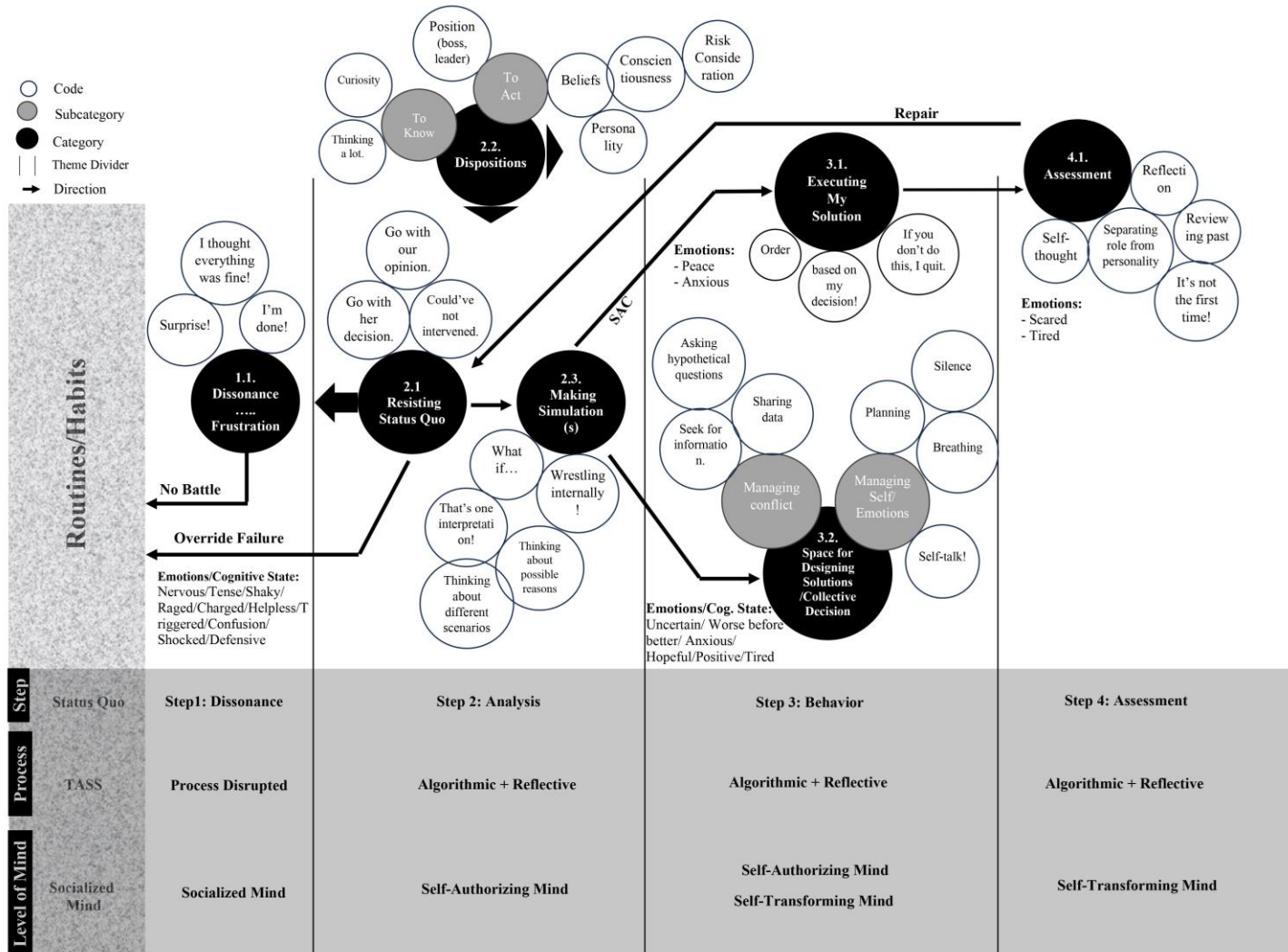
Concept Mapping

Inspired by Saldaña's (2021) operational model diagramming and Bhattacharya's (2017) information processing, a concept map/diagram of data was drawn showing the themes, categories, and subcategories that emerged as well as created codes and subcodes. According to Friese (2012), the aim of using concept maps in the form of network views for conceptual analysis is to "illustrate the answers to research questions instead of just displaying the categories. Such network views are likely to contain subcodes of a variety of categories, but only those that are relevant in the context of the research question" (p. 214). Figure 4.1. shows four big sections representing four main themes separated by three vertical black solid lines, seven big black circles depicting categories, four medium-sized gray circles as subcategories,

and 32 small circles representing codes. Each theme shows a unique mental step in the reasoning process and consists of a minimum of one category. The categories and their associated subcategories and codes are in touch with each other. There are also arrows between the sections that help with tracking the process and order of themes happening from left to right as the decision-making process addressing the qualitative research question. The hatched box before the first step of the process represents the status quo condition.

In response to “Why were circles and arrows used?”, the researcher assumes it comes from his interest in geometry and specifically graph theory. Also, perhaps part of it is because he identifies himself as a process-following person and tries to organize concepts in an associative manner. During the analysis, the researcher has been constantly reminding himself about this subjectivity that could contaminate the analysis and tried to bracket himself while doing the analysis. However, he does not claim that he was completely successful in doing this.

Figure 4.1. Concept Map



Theme 1: Participants' Decision-Making Process Starts After Dissonance

This theme describes the first mental step where a person's current situation is disrupted. This theme has only one category. Due to its conceptual difference and importance in initiating a reasoning process for a new behavior, the researcher decided to label this mental step as a distinct theme.

Category 1.1. Dissonance to Frustration

In response to the interviewer's question about describing a recent failed or successful leadership action, participants started sharing their story by referring to a surprising moment that disrupted the current established expectations, knowledge, opinions, and routines. Here, we name this disruption as cognitive dissonance and the current situation as status quo. These moments usually put the participants in a negative emotional state as they reported anxiety, uncertainty, tension, nervousness, defensiveness, epiphany, and shakiness. For instance, in describing her successful leadership action, Ashley, a community theatre producer, shared how she was surprised when a director decided to take an experienced actor's role from him,

We all have a different perception of what's a strength. And I did think I was a little surprised that she kind of came back pretty strongly and said, "I don't really want him in this role". And I thought, "Well, that's really surprising to me because I thought he'd be terrific". And so. I do think... some of the thought process was "I literally don't understand why you feel so strongly about him not being in this role".

Also, in her failed action description, this surprise happened differently for Ashley in dialogue with another colleague,

It was probably right before we started rehearsals. [My colleague] had said, "You have asked me to lead and sometimes I feel like you don't let me lead". I definitely felt

surprised that felt like new information. Also, in my mind, it felt like a pretty routine... I thought everything was fine and hmm, sounds like not. So, I do think that moment felt kind of shaky and uncertain. I thought ... this is different.

Another participant, Kait, who is an accessibility board member, describes her emotions as below when she realized the board is recruiting a new member with whom she disagrees,

I was aware of this person's work with disability simulations... I was absolutely just filled with rage. You know, so I had shared why I thought this was problematic.... I could feel my anxiety was up there and I could. It was a very emotionally charged moment for me because I knew I was taking a very strong stand.

There were many other stories about such difficult and even painful cognitive dissonance moments. Julia, an academic affairs coordinator, also shared her surprise when their team's progress was not acknowledged, and even besmirched in a meeting with top authority of the institution. Kerry, a health coordinator, described her feeling as appalled when she noticed their member visits could be unsafe.

Besides, there were stories illustrating that unaddressed dissonance after a while turned into frustration that eventually led to a new leadership action. For instance, Sydney, a non-profit director shared a story about how their enrollment rate and their community connections after the pandemic suddenly decreased. In response, she came up with an experiment in which her colleagues were not self-driven to participate. Not being able to address the problem, she became frustrated and after a while started a new leadership action. As part of this new effort, she convened a whole organization meeting where she realized how her knowledge about the problem was dissonant with her colleagues,

So, at one point during the meeting, someone said, “We are in a lot of the [communities]” I know the data and I know that not to be true and so like having to manage my emotions a little bit in that moment, ... I don’t want to say frustration. It wasn’t anger, but it’s like, “Come on.”

There were also strategies shared in participants’ stories about how to manage the emotions after dissonance. For instance, Tamara, a student support services director, mentioned she managed her emotions after getting feedback from her colleagues by separating her role from her personality. Angie as a CEO decided to be quiet upon realizing her colleagues were pursuing a divergent strategy for the project, she had been working on for nine months, contrary to her initial understanding.

Theme 2: Dissonance Triggers the Analysis

Theme 2 describes the second mental step which is initiated right after facing cognitive dissonance if the status quo is resisted. This section describes how this phase might be completed, left incomplete, or skipped which ultimately leads to different types of actions. Theme 2 contains two main categories that are explained below.

Category 2.1: Analyzing the Situation Starts with Resisting the Status Quo

On some occasions, especially failed or non-actions, people dealt with dissonance/frustration by ignoring it, justifying it, or bearing with it at least for a while. For instance, referring to her silence in the team meeting with an authority Julia mentioned,

You know, not every session is gonna be fixed, and not every session is gonna be perfect. Sometimes I think that’s one of those that you just have to, I hate to say, but let it go in the gap and address it later.

However, regardless of time and when it can happen, data suggests exercising leadership requires accepting dissonance and analyzing it for running a new action. This moment of acceptance was highlighted in several interviews. For instance, regarding the previously mentioned disagreement with a director, Ashley said,

... I could just say, "okay, that's fine". And I don't, I think sometimes there's nothing wrong with saying, "Okay, you have an opinion. Let's go with that."

Ashley also described this moment of ambivalence to act or not as *wrestling internally about the disagreement [with director]*. There were also other cases referring to this moment, for example, Kait referred to occasions when she would walk away instead of intervening.

Tamara shared a success story about a time when she received her colleagues' evaluation of her performance in the work environment. She described her reaction to a few critical feedback by emphasizing her self-questioning style,

Making me feel good as a director is not the priority. And sometimes it's hard to...I think I'm a pretty damn good boss. I'm pretty nice. I'm pretty easy going... You get all that kind of thing. But I don't think that's a good leader... I struggled with myself a little bit. I'm like, okay, did you fail in this because did you not approach those topics or was I just able to let it go because that is not what was important in the moment... Like was that the right decision in that moment?...

Participants referred to these moments as times when they were struggling to decide whether to acknowledge that the new situation requires a new response/behavior or not. We would call this mental step as the battle between status quo/habits/routines and an unfamiliar pull of new.

Category 2.2: Dispositions

What causes people to resist the status quo? In response to this question “what motivates you to act?”, data suggests there are several reasons why people resist the status quo and react to the dissonance and proceed with actions to change the current situation. Based on the interviewees’ responses, these motivations can be subcategorized into two groups. Some of these motivations are related to action such as role, job description, position, values, conscientiousness, risks, and beliefs, and some of them are related to their attitudes, and personalities such as the tendency to think a lot or to learn more. While the latter might be mostly related to making a new interpretation of the situation, the former seems to be more related to action and implementation. For instance, Ashley referred to the tendency to learn about the right answer while facing the dissonance and highlighted the value of collaborative work and her role as a leader while resolving the disagreement with the director. Sydney spoke of her belief in their work with vulnerable people as a motivation to find alternative better ways of solving problem.

In another case, Tamara interestingly mentioned several motivations together about why she accepted her colleagues’ critical feedback,

Really want to be the best leader that I can. It’s all of the things that you read, and my husband always says, “ask good questions”. Then, I think you know some of this is innate because I think some people just have a more reflective personality than others.... I am a lifelong learner. I always feel like I can be better. It doesn't matter what role.

Kait is so strong about her stand not to recruit the new board member that says,
Because there are some things where it is you do have to take that strong stand and that's one of them. You know, I think if you’re going to do that though, of course you have to

bear in mind that you can't be a snotty prima donna about it either and use that all the time.... You know, this is a hill I will die on, bring me my sword.

In contrast, regarding her own needs in the work environment, she explained that she refrained from intervening differently despite having potential solutions in mind. She shared that it might be because of what she learned from her mother in childhood about focusing on her abilities, not needs.

Moreover, Mary, a local business owner who shared about the local business problems in her community, in response to “Why did not you walk away and let it go?”, told a story about a tragic accident in her life when she lost her husband and two sons, and that changed her approach to impacting others’ lives.

Julia also shared a different type of disposition in mind in which she considered the risk of intervening. Her response suggests that she is still grappling with the decision of whether it would have been better to intervene or not,

...that was one of the times that I felt like I should have spoken, but I didn't. You know how you read the room.... No matter what was said or done. It felt like didn't matter. But on the hindsight, I felt like I should have come off the back and you still spoke anyway...But this was a different type of environment at that moment, especially with the prioritization going on. [It] was not the time to rock the boat...That is a don't do at that time you just don't well you know hey and I mean they're gonna be like, no, no, you need to go away.

Risk consideration was also repeated as a motivator of action (raising the heat) in Kerry's intervention as a health coordinator. Describing occasions in which they may not feel safe while visiting a member's home (due to their history), she said,

...That is one of my prime concerns and I think it's also a driver with why we are losing some of our staff is we are faced with some very uncertain safety risk...

While the risk around unsafety prompted Kerry to raise the issue with the authority, her colleagues had decided not to intervene, deeming it risky to their job,

I've had a lot of people like personally think me for bringing this issue up and what I have found out is people are just afraid to bring things up.

It is worth mentioning that Kerry also refers to the tendency to assess different scenarios before intervention. A tendency that was built in her following her past failed actions. This will be discussed further in the next section.

Data suggests the way these dispositions get involved in actions sets adaptive leadership interventions apart from other approaches to action. We assume that adaptive leadership actions entail the tendency to consider multiple aspects of the situation (such as learning and critical thinking, analyzing different scenarios, etc.), coupled with action-oriented dispositions (such as beliefs, roles, positions, values, etc.).

Category 2.3: Analyzing the Situation

After resisting the inertia to intervene, participants described a moment when they started to think about the dissonance/frustration/problem in a new way than before. In cases with a more adaptive leadership approach, this new way involves examining the problem through the lens of others. For instance, after a few months of observing her colleagues not being motivated to do what she asked them to do to solve the enrollment problem, Sydney wondered,

I think my realization over the past few months is not everybody knows what I know about the data of our organization and it's not fair to hold them to standards for data they don't know...

This new analysis or interpretation of the challenge led her to a new leadership action that will be discussed later.

A similar realization happened to Ashley when she came up with questions about the director's decision to learn more about the situation,

...my biggest action was asking hypothetical questions of "What is it look like if we cast him? What does it look like if we don't cast him, and we cast this other person?"

Kerry also referred to this type of analysis as going through scenarios,

I will admit that anytime that I plan to move something forward, even if it's just like a little planning, a little seed I want to plant, I go through every scenario...

Sometimes analyzing the situation may not lead to a successful adaptive leadership action or a leadership action at all perhaps due to not involving the lens of others. For instance, in her failed action as the committee meeting organizer, Sydney described how she decided to go as usual about the problem and finally quit the role even though she knew the causes,

We were working on team goals, and it was clear our experiment was failing. My reaction was, okay, that didn't work. I'm done trying.

Data also indicates instances where individuals made different interpretations or sought different perspectives of the problem but ultimately intervened based on their initial interpretation. For example, in her aforementioned failed action, Ashley responded to her colleague's surprising comment based on her own interpretation, despite attempting to understand her colleague's interpretation of the problem,

... [I got] stuck with one interpretation without the exploration of multiple.... in the moment I asked a few questions. Which I think is a good action. And then I made my own decision about what was needed. And I think the way it ended up kind of we had to like

muddle through a little bit. It was particularly tense for a little bit. Because I kind of jumped to a conclusion.

There were also cases such as Kait who approached with an interpretation of the dissonance and intervened without mentioning other alternatives or asking for other interpretations,

Well, so my thought process was that this person could steer conversations about accessibility into a way that's actually very ableist and into a way that does not actually look at the needs of disabled people.

Our observation of several stories suggests that, generally, the failed actions reported by participants were grounded in preferred single personal interpretations. This will be discussed further in the following section.

Theme 3: Participants Exercise Leadership by Creating a Communicative Space to Design a Solution While Exercise Authority by Executing Their Own Preferred Solution.

Comprised of three main categories, this theme delves into the third mental step which drives an individual toward action. Two different intervention approaches emerged in the coding process discussed below as separate categories. Although Category 2.2. was discussed under Theme 2, it is also effective on Mental step 3 especially through “dispositions to act”.

Category 3.1: Executing My Solution

This category represents the actions that people did based on their preferred interpretation. In these actions, nobody else was involved except the interviewees themselves. Sydney’s first unsuccessful attempt to increase the enrollment rate resembles a direction-order action,

... I tried to do this, "Okay, we're gonna set standards, and every month everybody is gonna hit at least 10 places". But it was really nebulous and it needed people to be very self-driven and creative. And not everybody here is super self-driven like that. Sometimes people need a task that's tangible. And so that experiment I tried back in March of like, "Here we're gonna do 10 a month blah blah blah to try to get these numbers up" wasn't working. We people weren't even getting 10 a month. We weren't seeing any change in our numbers.

Another action was Ashley's response to a colleague's critique that she evaluated as a failed action,

I asked a follow-up question and said, "Wait! When have I asked you to lead and then not let you?" ... I think it's the pieces that came next. That really probably gave me the bigger indicator. Because then we got into rehearsals and things felt tense between us....

Tamara described her immediate response to the conflict with a young colleague as below,

...It's just like, I'm done... I think I talked through my head like even previously like knowing that there might be confrontation and because there had been other things [happened before] I had already talked to my boss about. So, my exact reaction was pretty much that right like "you have to decide if you're gonna work with me and for me or we're done". And then when she just stared at me and didn't give me a response, I was like, "okay, I'm gonna go to HR..."

Another action was described by Kait who was even prepared to give up her position on the board if her opinion was not accepted,

Well, so in that specific moment, what I did was say, you know, "I object to this person

being on the board for”. So, and then. I'm very fortunate that I didn't, you know, the board respects my voice on these issues. And, in fact, one of my colleagues on the board said, “What I'm hearing here is a boundary”. Because I did flat-out and say that I would not serve on the board with this person.

These examples resemble direct interventions as definitive solutions by individuals. It reveals a type of authority dynamics wherein others are seen as the followers of the intervener. This dynamic is distinctively different from the type of intervention described in the following section.

Category 3.2: Creating Space for Designing Solutions.

This category mostly posits successful actions that do not necessarily execute the final solution for the problem but provide a space for others to collectively find a solution. Another feature of these actions is that they actively involve others' perspectives before taking action. Data suggests this type of action happens when participants realize there might be other perspectives about the problem at hand. This realization might come out of past failed individual actions or a general tendency toward collaborative work.

This new approach to solving the problem typically put people in situations where they frequently reported the need to manage emotions (internally) and conflict (externally). In response to the question about their feelings during the interventions, people still reported nervousness, anxiety, or worse before getting better. For instance, Ashley described her successful intervention in her disagreement with the director as below,

Okay, I would say, “well, here's what I see” and another team member would say, “that's one interpretation”. Yeah. and, then the third one would say, “Okay, well, another interpretation is”, ...And I think it was really great for us to be able to say,

there's a multitude of good solutions, you know.... And then I think we landed on a casting choice with intentional [conversation] I think. And then, ended up being good for the whole show...I think we were able to kind of yeah probably work through it well. And that felt successful.

Asking about the emotions during that intervention, she provided more details about what happened in the conversation and how she managed the conflict and her emotions,

I mean, on one hand, I don't know, like it gets worse before it gets better kind of feeling you know cause she said, "No! I don't want him there", and then to say, "Okay, Can I ask why you feel that way? and can I ask where the answers coming from?" And then for her to say, "You know, he was kind of derailing and how he was acting in the last show". And that feels like a legitimate concern to say, well, we need someone who works well with others and is working towards the main show. And so, I probably felt more nervous to say, "That's a real reason why not to cast someone?" And then the other person said, "Well, yeah, that's true. But their wealth of character development really enriches." ... [And to manage my emotions] I think in the moment, lots of breathing, some even like inner self-talk and suspending judgment.

Here, the participant is describing how she benefited from asking multiple questions and inviting others' perspectives to manage the conflict.

Sydney also shared their group meeting in response to the low enrollment rate problem as a successful intervention,

I used my locality workbook a lot in drafting my agenda for that meeting. And I send them out to everybody in advance, so they know what questions I'm going to ask. They can come prepared. I think that's worked out well. We've had conversations that we

traditionally have not had in the organization.

Again, this participant surfaced the ways she managed conflict and emotions during that session,

At one point during the meeting, someone said, we are in a lot of the [communities]. It's kind of like, we're already doing this. And I was like I know the data by an organization and I know that not to be true and so like having to manage my emotions a little bit in that moment, but also make sure the group didn't go down that path, cause that's just not true. And so that was kind of one of those moments of like raise the heat moment. [Then, I said,] "Well, let's look at that. Let's pull up the database and let's see. There are 56 [target areas] in the district.... We're in three of them"

Here, the participant describes how she managed the disagreement by sharing more information and regulating her emotions to keep the conversation going. Additionally, when asked how she managed her frustration during the session, she said, *"by planning"* and *"getting help from advocates in the meeting"*.

This type of meeting, filled with conflicts and charged emotions, to reach a consensus on the experiment was repeated as a successful action by several other participants. Mary also referred to a collective meeting to solve the grocery distribution issue in her community. Angie, Kerry, and Julia also reported the initiation of such collective dialogues that need time for their outcomes to become evident.

Theme 4: Participants May Initiate a Repairing Process After Failure

Category 4.1: Assessment/Reflection

There were a few participants who referred to a moment in which they assessed their failed actions and started a new intervention. After describing the failed actions, they shared a self-reflection process that, in some cases, was initiated by their dispositions, while in some cases, it was influenced by fatigue, leading to a new action. Their description of the assessment shows that in those moments they went back to mental Step 2 and tried to come up with a new simulation for the problem.

After describing her failed intervention in her interaction with a young colleague, Tamara shared her reflection process, which took a while for her and led to a new conversation with the colleague. When asked what made you reflect on that experience and go back and revise your behavior, she shared multiple reasons. First, this reflection started when she noticed the same conflict happened on other occasions and made her think about the problem more and deeper. Second, her upbringing as the youngest child made her observe others and learn through experiences. Third, the value of teamwork. Fourth, she frequently refers to an innate motivator that can't be clearly described,

Why do I reflect a lot? I always reflected as a kid or even in college. I don't know. All I know is I've always wanted to do things better. So, I don't know if it's that innate personality trait where if I want to do things better....

Sydney also shared such a reflection moment (described earlier) where she tried to think about the problem from others' lenses after initial failed action. Additionally, Ashley, referring to the value of team and relationships as motivators, later invited her colleague to a coffee conversation to talk about the conflict. She said,

I think if I pressed in and asked more questions. I think it could have been a more successful experience of leadership.

Angie also referred to a meeting with colleagues where she realized her colleagues were not considering her approach to a project despite her efforts over months. She assessed her leadership as failed since in her reflection time (over a weekend) she noticed she had not been listening to her colleagues at all. Moreover, like Tamara, she realized that this was not the only time she failed by not listening. Angie started a new intervention after this reflection time which she calls it “Repair”,

I said, “okay, I’ve got to do something. I’ve got to repair this. I’ve got to build bridges. I need to figure out how we can then now all work together”. So, then I sent a couple of texts like, “hey, thank you so much for saying something in the meeting. I appreciate these ideas and I’m sure that this will work out. Let’s just all get back on board. So, there’s some like repair the following week”.

She reported her feeling in this new action as “scared”. In response to how she controlled her emotions, she said,

I go and talk to somebody who just loves me or thinks the world of me [to] get all those pats on the back and then I’m like, huh. Okay, then I can face this scary thing, you know, or a relationship where I think the person on the other side is really not happy with me. So, I guess I boost my ego. And then I can go forward.

Reviewing participants stories about the new intervention after the reflection time, it seems besides thinking dispositions, their realization of a pattern of failure is also in place as a motivator. This was evident in the response of Tim as an organizational director, who even reported experiencing fatigue due to repeating the same failed behavior. Overall, it seems like

failed experience may come as a new dissonance or frustration to individuals that again push them back to the mental steps 2 and 3 to design a new intervention after acknowledging others' perspectives.

Discussion of the Findings

Discussion of the Validated Scales

Adaptive Leadership Behavior Scale (ALB): Leadership Requires Surfacing Conflict

The four factors of ALB represent a type of leadership behavior that resides at Kegan's self-transforming mind with the highest level of mental complexity. According to Kegan and Lahey (2009), at this level, the mind stands back from its own filter and make space for modification of design or agenda, seek for information from others, manage optimal conflict, and sees its theory, framework, or ideology as drafts not finished products. These all align with the four factors emerged in the ALB validation process: 1) seek stakeholders' perspective 2) Surface conflict 3) Make space for multiple perspectives 4) Act experimentally. This receives more importance as we see Kegan and Lahey (2009) argued that solving adaptive challenges requires self-transforming mind as a sophisticated stage of mental development. It can be concluded that the higher ALB means the higher mental complexity. This opens the stage for exploring the relationship between ALB and mental complexity as part of the third hypothesis in this research.

Another important aspect of this scale is the emergence of "conflict" as a factor. Besides Kegan and Lahey's (2009) emphasis on "managing conflict" as a feature of complex mind, conflict has been identified as an important concept in adaptive leadership literature. According to Heifetz (1994), "The exposure and orchestration of conflict – internal contradictions – within individuals and constituencies provide the leverage for mobilizing people to learn new ways" (p. 22). Heifetz et al. (2009) also asserted that "...to practice leadership, you need to accept that you are in the business of generating chaos, confusion, and conflict, for yourself and others around you" (p. 206). Although mentioned in Chapter 2, it is worth repeating that Avolio (1999), referring to Heifetz (1994), considered adaptive conflict as the basis for building a highly

developed team leadership. This means that ALB can be used to assess people's capacity to build effective teams, especially in companies and organizations.

One more important point about ALB items is the emphasis on the two-way communicative essence of leadership behavior rather than solely transferring or influencing actions. The communication style in this type of leadership is transactional in contrast with more traditional approaches to leadership in which communication is employed as a transferring or strategic tool to inform or influence the followers. According to Eisenberg et al. (2017), "unlike in transactional-process model, which assumes that effective communicators are clear and open in their efforts to promote understanding and shared meaning, the strategic-control perspective regards communication as a tool for influencing and shaping one's environment" (p. 236). ALB items, especially factors 2 and 3, describe actions not as finished ultimate solutions but as an invitation message that a person sends to engage others in a shared understanding of the challenge.

Heifetz (1994) reflects on this distinction by inviting the readers to make differences between these two sentences "Leadership means influencing the community to follow the leader's vision" versus "leadership means influencing the community to face its problems" (p. 14). ALB highlights actions that need communication skills to build a shared space to find the solution regardless of who is leader or follower. The first three factors (seeking perspectives, making space, surface conflict) resonate with Heifetz's quote in his 1994 book: "the inclusion of competing value perspectives may be essential to adaptive success" (p. 23). ALB emphasizes the two-way relational behavior of an individual, wherein they engage others not only to seek perspectives and get feedback on their ideas but also discuss one's own commitments.

Moreover, Heifetz (1994) distinguishes adaptive leadership from transformational leadership as another relational and transactional approach to leadership by emphasizing their differences regarding purpose. According to Heifetz (1994), the measure in adaptive leadership is progress on problems, not influence. We believe that ALB shows the behaviors that are required for progress on the challenge not influencing others particularly through Factor 4 (Act experimentally). The actions of other factors also revolve around the concept of mobilization not influencing others (Heifetz, 1994). This may raise the question whether adaptive leadership behavior should be considered as an outcome variable especially in relationship to leaders/ship identity or a precursor.

In Bass's (1985) conceptualization, transformational leadership aims to stimulate followers to work on the leader's inspiring vision. It seems the solutions and goals are available and influencing followers and transforming them is the work needed to be done by the leader. In adaptive leadership, solution is not a reality independent of mind in the hands of a leader. It is something that should be constructed through engaging multiple perspectives and tested through subsequent multiple experiments. Hence, the goal is not to motivate others to work toward a defined goal. The goal is the progress on the challenge that necessitates mobilization of others.

Adaptive Leadership Self-Efficacy (ALSE)

The two-factor ALSE scale represents the "confidence to diagnose and energize others". These factors and items resonate with Hannah et al.'s (2008) suggestion that leader efficacy should incorporate both actions and capability to interpret the context and generate novel solutions. While Factor 1 presents items related to the capability of challenge analysis and designing solutions, Factor 2 contains items of action and execution.

Adaptive Leadership Self-Awareness (ALSA)

This scale captures the first component of Taylor's (2010) self-awareness which is about awareness of self-character, traits, assessment of strengths and weaknesses, sense of purpose or calling, core values, and motivations and desires. It also speaks to Ferrari and Sternberg's (1998) awareness of other selves particularly through items ALSA9 and ALSA10.

ALSA's second factor, *awareness of role and purpose*, is an important component of knowing self as a system in Heifetz's work that emerged in the ALSA validation process. According to Heifetz et al. (2009), knowing your role provides you with emotional strength against personal attacks that may impede interventions. Moreover, highlighting an important aspect of the reflective mind, Heifetz et al. (2009) referred to "knowing your purpose" as part of seeing oneself as a system. They argue that purpose plays the role of director to action and reminder of the reason when things get difficult. This was also reverberated in Day et al. (2009) as goal orientation.

Another aspect of self-awareness in Heifetz's views is understanding defaults. This component appeared in two separate factors "awareness of defaults" and "awareness of needs". In this domain, while knowing loyalties did not remain among the final items in the validation process, items related to bandwidth (comfort zone) and personal tuning were able to make to the final version of the scale. Heifetz et al. (2009) asserted that knowing these aspects of self helps you with more proactive and efficient interventions as you step into uncertainty.

Putting this all together, the factors that emerged in ALSA, though without the "knowing loyalties" component, are expected to contribute to more adaptive leadership behavior. This will be explored further later in this research as part of the third hypothesis.

Adaptive Leadership Shifting-Reflectiveness (ALSR)

Developing this scale, though not followed the best practices like other three, aimed to explore a higher-order cognitive process in decision-making toward adaptive intervention. Informed by the third cognitive decoupling in Stanovich's (2008, 2011) tripartite model and Heifetz's "will" as a lay psychology term, the researcher aimed to see if a higher-order thinking disposition – distinct from MRM – can be captured within the reasoning process. This higher-order disposition is assumed to be a driver toward adaptive intervention. This hypothesis was confirmed in this research, and the two-factor scale gives insights about the important drivers of actions that Heifetz highlights in exercising leadership.

The importance of the first factor, *purposefulness*, in adaptive leadership was somehow discussed earlier in the ALSA discussion. According to Heifetz et al. (2009), "There is no reason to shoulder the difficult work of leadership if you do not have compelling, higher purposes to serve" (p. 233). They invite people to list the purposes underly the adaptive change they are seeking. While people may have different personal and contextual purposes to exercise adaptive action, we tried to reflect on them through generalized statements. However, we acknowledge that these statements may not encompass all the thinking dispositions/purposes needed for intervention. Further exploration of dispositions related to shifting reflectiveness in adaptive actions is recommended.

The second factor underscores the distinguishing feature of adaptive leadership, namely its departure from traditional authority-based approaches. This factor underscores the mindset of not relying on authority as a driver of adaptive action. In other words, we propose that ALSR4, ALSR5, and ALSR7 can serve as higher-order motivators of action. We posit that one way to distinguish adaptive leadership practitioners from others is through this mindset. While

individuals may generate various strategies for addressing a problem, some may defer action to others, particularly authorities. This mindset serves as a link between Cognitive Decoupling 3 and action.

From a different perspective, the two factors are separating two different loci of control in human cognition. While the first factor is more emphasizing the internal locus of control, the second factor is reflective of the external locus of control. According to Bulus (2011), people with lower internal and higher external loci have a lower mastery goal orientation and higher avoidant goal orientation. This perhaps suggests that ALSR has the capacity to reveal the work avoidance dispositions.

Model Choice Discussion

LMM10, supports our hypothesis regarding the relationship between adaptive leadership behavior and reflective mind components, explaining 35% of the variation in adaptive leadership behavior as indicated by the R-squared marginal. It is worth mentioning that, including TLB instead of ALB, in the model reduced the model fit significantly and resulted in an inefficient model.

Model 10 gives more details about this unique relationship accounting for individual differences in the relationship between ALB and MRM. MRM's fixed effect, interaction with ALSE, and the interaction's random slopes significantly predict ALB. In this research, MRM was considered as an assessment tool to measure the tendency for decoupling and, therefore, complex thinking. MRM individually predicts about 9% of ALB. It predicts 3.4% of the variation in ALB after controlling for other reflective mind components and demographic criteria. This confirms that MRM can be used as a tool to map people's cognitive complexity on Kegan and Lahey's (2009) mental complexity model (socialized, self-authoring, and self-

transforming minds) and the adaptation tool (Figure 2.10) developed in this research in Chapter 2. The higher the MRM, the more tendency an individual might possess in making metarepresentations and the more adaptive leadership behavior they exhibit.

The random slope of MRM's interaction with ALSE have high negative correlation with random intercept. This suggests that at higher leadership baselines the association between ALSE and ALB is weaker. This is probably due to ceiling effect or saturation. Additionally, this correlation might give more insights about the individual differences that predict about 40% of ALB ($R^2(C) - R^2(M)$). We do not have any idea at the moment about these individual differences. However, we reframe this result here in a way that it may assist future researchers in investigating it: "There are individual differences that influence your leadership behavior baseline that are negatively correlated with individual differences that influence the ALSE*MRM."

ALSA individually predicts about 18% of variance in ALB and 3.4% (R-squared Marginal) after controlling for other reflective mind components and individual differences. Moreover, among all IVs, ALSR showed the highest explanatory power predicting about six percent of unique variance in ALB while controlling other variables and individual differences. As discussed in Chapter 2, shifting reflectiveness accounts for driver dispositions toward action. This shows the importance of cognitive decoupling 3 and highlights the importance of developing a more quality ALSR scale in the future.

Signs of Algorithmic Mind

MRM's significant random slope in the model, though with small effect, suggests that the relationship between cognitive decoupling 1- 2 (overriding the status quo and making multiple simulations) and ALB is not consistent among individuals. For instance, when two individuals

make N interpretations about a situation, this does not necessarily result in an equal number of leadership actions being taken by each individual. We assume that due to the entanglement between the reflective mind and the algorithmic mind in the reasoning process, these individual differences might be related to differences in the cognitive capacity for running algorithmic processes or phenomena related to the taxonomy of thinking errors. There is also another hypothesis that highlights the role of cognitive dissonance – as the initiator of the reasoning process for new behavior – in defining those individual differences. This will be discussed further in the qualitative discussion.

The random slope of MRM's interaction with ALSE successfully loaded in the model serves as additional support for our theoretical framework. The presence of this interaction's random slope, coupled with individual differences reflected in the intercept, elucidates that the 40% variation captured by random effects extends to other cognitive differences influencing the dynamic between MRM, ALSE, and ALB. While acknowledging the potential existence of other reflective mind components tied to cognitive decoupling, our hypothesis posits the significance of the algorithmic mind in shaping leadership actions. We posit that a substantial portion of the unexplained 40% variation in random effects can be attributed to the algorithmic mind processes.

The interaction between MRM and ALSE highlights a nuanced connection between cognitive processes and adaptive leadership behavior. This interaction indicates that the impact of MRM on ALB is contingent upon ALSE, and vice versa. Put simply, the relationship between individuals' MRM and ALB strengthens as levels of ALSE increase. This observation gains significance when considering self-efficacy as a component of Stanovich's fluid rationality, which encompasses computational capacity. If so, within this framework, LMM10 proposes a model wherein both algorithmic and reflective cognitive processes serve as predictors of

adaptive leadership behavior. Furthermore, this model illustrates an interaction between these two cognitive processes. In Heifetz's terms, this suggests that while skill and will individually lead to leadership intervention, skills become more influential at higher levels of will or purpose.

Discussion of Path 2023 Impact

The evaluation study examined changes in cognitive (ALSA, ALSR, MRM, and ALSE) and behavioral (ALB) outcomes. Data suggests that ALB increases significantly at T3 and 4. This shows the impact of the program on adaptive leadership behavior. Moreover, the program improved all reflective mind components except for ALSE. It can be argued that the program had a positive impact on people's tendency to make multiple interpretations (MRM), awareness of defaults, needs, and emotions (ALSA), and action orientation (ALSR). The smallest increase among all was observed in MRM which could be expected as it reflects on the first burdensome driving cognitive function for adaptive leadership action that is responsible for disrupting the status quo and making a new representational model for the situation.

LFC evaluation results also show improvement in adaptive leadership competencies three months after the program. However, comparing the group of people who attended the program individually and those in their organizational groups, it seems like the program is more effective on individuals. There might be several interpretations behind this that require further research. One interpretation is that participating in LFC with your organizational group may not necessarily establish a commitment to intervene. People in groups may struggle with defining whose work it is to intervene as a result of their discussion in their groups and they wait for others to take action.

The successfully converged model provides detailed insights into the relationships between reflective mind components and ALB. Notably, the effect of ALSA on ALB

demonstrates a positive temporal shift, a pattern observed similarly for ALSR and ALSE. These changes are associated with a training program implemented between T1 and T2. While a more confident claim about this effect awaits experimental validation, it is assumed that over time, the effects of ALSA, ALSR, and ALSE on ALB will increase. Attributing Time to the YLE program here, we argue that the program's effect on people's ALB is higher at higher levels of ALSA, ALSE, and ALSR. In other words, people with lower ALSA may not see the same increase in their ALB as people with higher ALSA. This dynamic between ALSR, ALSE, and ALSA and Time does not repeat for MRM. The program's effect on people's leadership behavior is consistent across different levels of MRM.

Moreover, evaluation results showed that the increase in ALSE in KLC's path was not significant even four months after the program. One potential interpretation is that unlike other measurements in this research, self-efficacy is considered as a positive psychological resource. According to the literature review in Chapter 2, adaptive leadership mostly benefits from negative affective resources. This is probably embedded in the teaching framework that the KLC team does not focus on individuals' self-efficacy. Considering its individual effect on ALB as well as its moderation effect on the relationship between MRM and ALB, it is recommended that ALSE receives more attention in the path.

Discussion of the Qualitative Phase

It's the Pain that Leads Human (Rumi, 13AD)

The qualitative interviews provided substantial support for our theoretical framework regarding the decision-making process. The decision-making steps described in the qualitative data representation is in alignment with the reasoning process for rational action hypothesized by Stanovich (2008, 2011). Individuals initiate the decision-making process for taking leadership action by disrupting the status quo. This aligns with the first step in Stanovich's Cognitive Decoupling 1 which is about inhibiting TASS. This may happen soon after a cognitive dissonance occurs or after a while that dissonance turns into frustration due to not responding to it.

This speaks to Levy et al. (2018) that argued cognitive inconsistency may evoke a negative affective state. According to Festinger (1959), this inconsistency among individuals' beliefs and behaviors causes psychological tension that may lead to a positive change in action. According to Krcmar et al. (2019), "in this situation, a person engages in a behavior that is dissonant with an attitude in an attempt to change the ratio of dissonant behaviors to consonant behaviors" (p. 219). This approach to reducing the dissonance ratio $\left(\frac{\text{Dissonant behaviors}}{\text{Consonant behaviors}}\right)$ is called "attitude bolstering" by Sherman and Gorkin (1980). Here, we call turning dissonant behaviors into consonant to reduce the ratio as "Leadership". Krcmar et al. (2019) also introduced another way of reducing this painful dissonance which is altering the importance of dissonant beliefs. We label this as Heifetz's "work avoidance" and discuss some of the cognitive mechanisms here using Stanovich's (2011) thinking errors.

The determination to disrupt the status quo hinges on the thinking dispositions that act as the commander of adaptive leadership behavior. We do not call dispositions' involvement as a step here as it is happening in parallel with other steps run by the algorithmic mind. Evans and Stanovich (2013) called these dispositions as default interventionist meaning that "Type 1 processing generates intuitive default responses on which subsequent reflective processing (Type 2) may or may not intervene" (p. 227). Within the category of dispositions, individuals shared various motivations, aligning with Stanovich's thinking dispositions list, including consideration of future consequences (risk), curiosity, and need for cognition (i.e., tendency to think a lot; Cacioppo et al. (1996)). Besides, there were common dispositions observed among both adaptive and non-adaptive leadership actors, primarily centered around the execution of actions, encompassing epistemic values, beliefs identification, role, and job description which may play the role of mastery or performance goal orientations. In this research, these two types of dispositions were measured separately in the quantitative phase via MRM and ALSR scales. However, studying the extent to which dispositions to actions and the tendencies to make multiple interpretations are mutually exclusive remains a potential avenue for future research.

In the second step, interviews revealed that in their adaptive leadership actions participants try to come up with a new analysis of the situation. This aligns with Heifetz's term "new interpretation" and Stanovich's concept of Cognitive Decoupling 2 or updating reflectiveness. This involves updating the existing model and creating a representational model of the situation to adapt to the situation. This is the phase which participants reported questioning themselves, looking at the problem from different lenses, and reflecting as their internal practices.

Following the second step, in their successful actions, people reported engagement in a dialogic space where they seek perspective from others and try to understand the situation from different angles. The outcome of this third step would be a final solution/action on which different people worked together. In Stanovich's tripartite model, this phase is most similar to Cognitive Decoupling 3 as a cognitive process through which people sustain making multiple simulations, pick one or integrate them and put it into action. In a variety of successful adaptive leadership actions described, people created this space. However, despite Stanovich's model, this step does not happen merely in the individual's mind. Informed by internal new simulation, they develop a new relational space in which multiple minds would be involved to continue making simulations or interpretations. From Kegan's perspective, these individuals perhaps move to the most complex level of mind: Self-transforming mind. The successful or repairing actions taken by the non-profit director, student support services director, and theatre producer serve as prime illustrations of putting such decision-making processes into practice.

The emergence of reflection phase (Mental Step 4) in the reasoning process during the interviews opens a new avenue for further exploration. The researcher did not account for this phase in the quantitative research. However, it seems like this phase is supporting Kolb's circular experiential learning model. Participants' failed actions and what they learned in the reflection phase came as a new dissonance or frustration for them. Moderated by their dispositions, this dissonance/frustration triggered restarting the process by updating the representational model of the situation (Mental Step 2) which ultimately led to seeking others' perspectives. The student support services' director, the CEO, and the organizational director clearly mentioned this realization during their reflection time where they learned their failure had a history. This realization prompted them to disrupt this habit (TASS) and design a new intervention.

Thinking Errors as Doors to Exit the Adaptive Process

Those who did not complete the adaptive leadership decision-making process can be categorized into three groups:

Group 1: Among interviewees, there were individuals who implemented solutions to their problems without completing the entire process. Some even did not accomplish the initial disruption phase (Decoupling 1) and behaved based on their existing model available in TASS.

Group 2: Some were able to pass this phase but did not finish Decoupling 2 (making more simulations).

Group 3: There were cases in which participants tried to update their existing simulation of the situation, evident in their pursuit of other perspectives. However, they ultimately acted on their own preferred interpretation.

The individuals in the first group did not seek for others' perspectives which emanates from not updating their own interpretation/simulation. In other words, they behaved based on their Type 1 processes. Some of these individuals evaluated their actions as failed and some as successful. Descriptions revealed that the failure of these actions can be attributed to Stanovich and West's (2011) taxonomy of thinking errors such as ego-centric processing (e.g., *my approach is the best and I will die for it or I'm the boss*), contaminated mindware (e.g., *I tried this way before and it did not work*), and mindware gap (*I did not have enough information, so I did it as usual*). The board member's self-evaluated successful action, the student support services director's failed action, and CEO's failed action may fall under these groups of errors. Kegan probably would locate the mental complexity of these people between the socialized mind and self-authorizing levels of mind.

Stanovich would characterize the second group, those who could not finish the decoupling process (updating simulations), as trapped in override failure. This implies that they initiated the process of making multiple interpretations, but they could not finish it due to a variety of reasons such as my side bias or mindware gap. Illustratively, the board member's and non-profit director's failed actions are two noteworthy examples of this phenomenon.

The third group finished the second step. However, we would not call this an adaptive leadership behavior here, given its lack of engagement with other perspectives. In fact, this is similar to a thinking error highlighted by Stanovich named Serial Associative Cognition (SAC) where individuals might search for multiple interpretations/simulations but act on their self-preferred model. It is presumed that Kegan would assess the mental complexity of individuals exhibiting this behavior at the level of the self-authorizing mind. A pertinent example of such a reasoning process can be found in the theatre director's failed action.

Not Conforming, Deliberative, Relying on Diversity, and Time Consuming

The key words in the heading are the main elements that Heifetz (1994) used in his definition of adaptation (See operationalization of constructs in Chapter 1). These elements were reflected upon in the decision-making process described in the qualitative phase:

- Mental Step 1: Not conforming (to cognitive dissonance).
- Mental Step 2: Relying on diversity (of interpretations).
- Mental Step 3: Deliberative (essence of solution designing).
- Mental Step 4: While time-consuming is the main feature of the whole problem-solving process with an adaptive approach, it was emphasized by people when they described the reflection phase.

People's decision-making process illustrates how adaptation happens by addressing cognitive dissonance in a non-conforming way through incorporating other perspectives with a deliberative approach. Using the process described above as a model, adaptive leadership behavior can be distinguished from other types of leadership behavior that are reliant on authority contract or leader-follower dynamics. The pain that comes with dissonance is the initiator of intervention, not influence. The solution emerges out of multiple perspectives rather than individual's preferred action. Others are involved in the decision-making process rather than being merely a follower or observer.

Mixed Methods Discussion

There are three major reasons why the qualitative phase of this research was designed to be follow the quantitative phase. First, to explain the findings of the quantitative phase particularly the relationship between adaptive leadership behavior and the reflective mind components; Second, to address the surprises that may emerge during the quantitative phase; Third, to offer practical recommendations for the improvement of future LDPs for KLC's teaching team. All these three help the research provide more insights about the gap between learning and practicing leadership. Leaving the third one for Chapter 5, here, we discuss how the first two reasons/goals were accomplished by reflecting on the results presented in the previous section.

Qualitative Study Provides More Information about the Quantitative Results

Validated Scales. Although the qualitative research was not designed to corroborate the scale factors that emerged in the quantitative phase, it is worth comparing the factors emerged in the quantitative phase with the types of leadership activities shared by participants in the

interviews. This may reveal potential actions that might not be captured by the scale but could be added to the list in future studies.

Participants shared several leadership activities aligned with the ALB's four factors. Here are a few examples:

1) Seek stakeholders' perspectives: The community theatre producer created a space to seek the director and stage manager's perspectives before casting. The CEO started seeking perspectives after acknowledging her failed action.

2) Surface conflict: The non-profit director convened a non-traditional session to raise the heat and seek perspectives and new solutions. Both the non-profit director and community theatre producer managed the conflict by seeking perspectives and managing their own emotions. Health coordinator's intervention also was labeled as "raise the temperature" by herself.

3) Make space for multiple perspectives: The business owner convened a session with other business owners to discuss the distribution problems.

There were also other adaptive leadership actions shared by participants but were not included in the ALB scale's final list. For instance, the academic affair coordinator's risk assessment that led to her strategic silence in the meeting with authority could be considered an example of "intervening skillfully". Item ALB 12: "I assess the situation before heating up an issue" perfectly encapsulates this behavior; however, it was removed from the list after EFA. From a theoretical perspective, this removal can be justified as it is difficult to distinguish a strategic silence from a work-avoiding silence. Perhaps, this item's low factor loading in the EFA emanates from its low distinctive power.

Also, there were several overlaps between the thinking dispositions shared in the interviews and those listed in ALSR and MRM. Table 4.23 presents examples of such disposition codes and the particular items they allude to. Moreover, the researcher noticed some codes that could be included as a disposition item in ALSR in future research. For instance, there was a URM participant who mentioned her identity as a Hispanic woman motivates her to show leadership quality. This was also repeated differently in an interview with an African American woman. Although these codes may align with ALSR1 (personal priority to exercise leadership), incorporating additional items focused on race could enhance the scale's specificity and discernment. This is also in alignment with KLC's YLE 2021 report that says URM people show more leadership behavior than others. Referring to Sternberg (2004), Stanovich (2011) also raised this curiosity that what if race impacts rational action as it involves thinking dispositions?

There were several moments in participants' stories that highlight the importance of self-awareness in exercising leadership and point to ALSA items. For instance, Tamara, the student support services' director below describes how her level of tolerance of confrontation formed her action.

I mean to be perfectly honest like I knew going into it there might be confrontation just because of the past experience I had. I would definitely go into it nervous because I told you I do not like confrontation.... But obviously, I mean, we gotta do hard things... When it comes to something where I think it's going to be confrontational, I just, I kind of lose my mind and that is what I'm working on.

This quote resonates with awareness of defaults particularly items ALSA 6, 7, and 8. She also mentioned how she managed her emotions in this confrontation as if she is referring to ALSA9,

Trying to breathe through it. I tried to tell myself like none of this is personal...

Also, in response to what made you choose to intervene rather than follow the status quo, Ashley, the theater producer referred to her role which resonates with ALSA1 and said,

I think probably as the leader of the organization, I probably was feeling and thinking more about the community as a whole... And how will this person feel about not getting apart...

Factor 3 mainly appeared in the reflection phase (Mental Step 4), where people initiated a repairing process once they realized they need to change their behavior. As described earlier, this appeared in the CEO, theater producer, and the student services director's responses. ALSA5 (*I am aware of my unmet personal needs....*) was parroted frequently in participants' self-reported impediments of leadership behavior. For instance, it was echoed in the board member's reaction to obstacles in taking action, as she acknowledged an old and currently wrong mindset ingrained in her since childhood that continues to influence her behavior.

Table 4.23. Qualitative Codes that Resonate with MRM and ALSR Items

Scales	Items	Disposition Codes Per Item
	(See Stanovich, 2011, p. 88)	
MRM		MRM2. Board member: <i>"You have to bear in mind that you can't be a snooty prima donna about it and use that all the time. That has to be for something that is. You know, this is a hill I will die on, bring me my sword."</i> MRM 8, 14, and 15. Health Coordinator: <i>"Going through every scenario before action"</i>. (Consideration of future consequences) Business owner: <i>"I like to think a lot..."</i> (Need for cognition) MRM4. Theatre producer: <i>"I want to know the right answer..."</i> MRM12. Student support services: <i>"She is the director of the programming; you will do what she says. Unless it's unethical, if she tells you to print it on yellow paper, that's what you're gonna do."</i>
ALSR	ALSR 1. Creating change at work/in my community is my personal priority. ALSR2. Creating change is part of my job description. ALSR3. I practice leadership because a person should always consider new ways of solving a problem. ALSR4. I believe we should look people in positions of authority to solve problems. ALSR5. Difficulties can usually be overcome through waiting for good fortune, rather than thinking about the problem. ALSR7. I believe there is little I can do to change the way things are.	ALSR1. Educational center director: <i>"Helping kids means something to me"</i> (conscientiousness) (non-epistemic value) Student support services director: <i>"As a Hispanic woman, it's important to show leadership qualities ..."</i> (non-epistemic value) ALSR2. Educational center director: <i>"I really believe in what we do. And This is mission-driven, like it's specifically focused on our mission and getting our mission out in the community."</i> ALSR 4. Theatre producer: <i>"Technically, I am the boss. I am over her ..."</i> ALSR5. Theatre producer: <i>"I could just say, okay, that's fine. I think sometimes there's nothing wrong with saying, okay, you have an opinion. Let's go with that..."</i> Academic affairs coordinator: <i>"Sometimes I think that's one of those that you just have to, I hate to say, but let it go in the gap and address it later."</i>

Decision-Making Process. In the qualitative discussion, we showed how different components of adaption were illustrated in the interviews. Additionally, as discussed in chapters 1 and 2, there is another definition of adaptation that connects adaptive leadership to DPT: Will plus skill or algorithmic mind + Reflective mind. The qualitative phase also shed light on the decision-making process underlying adaptive leadership behavior. In the quantitative phase, it was hypothesized that there is a relationship between ALB and the reflective mind components (MRM, ALSA, ALSE, and ALSR). Examining how people describe their reasoning processes to exercise leadership showed that MRM as the tendency to decouple from the existing representational model of the situation can be found in people's "dispositions to know". Moreover, ALSR as the disposition to sustain continued simulation to act is evident in people's "disposition to act". Although data suggests any type of leadership action may involve the disposition to act, ALSR's second factor emphasizes its adaptive aspect and distinguishes it from other types of goal orientation. Our effort to validate ALSR here might be considered insufficient due to time and data collection constraints. However, its significant impact on ALB in the LMM model suggests that further exploration and strengthening of this measurement tool is warranted.

Interviews also revealed epistemic and non-epistemic values as dispositions that are involved in leadership action. Stanovich (2011) highlighted such values in his model. Bringing examples from the interviews, dispositions that have origins in "race" or "conscientiousness" can be categorized as non-epistemic values while "tendency to go through every scenario" and "think a lot" can be considered as epistemic values.

Heifetz's emphasis on knowing self as system, Day et al.'s (2009) integrative leadership development theory, and Stanovich's (2011) concept of controlled emotional regulation as part of reflective mind processes encouraged recognition and validation of ALSA as a new scale.

Moreover, Wrana (2014) argued that self-awareness is key in adaptive emotion regulation strategies. The identification of “surface conflict” as a factor in ALB and its associated items reinforced the importance of emotional regulation as an important act of leadership.

ALSA was validated as a new scale and was included in the adaptation model in the quantitative phase significantly predicting ALB. However, uncertainties persisted regarding how and when in the reasoning process ALSA will play its specific role. The interviews revealed several cases where people refer to emotional regulation which requires conscious awareness of self, particularly during the moment of cognitive dissonance and within action space. These steps include moments with negative emotions such as nervousness, anxiety, uncertainty, and shakiness where people’s interpretations are being jolted or collided. Individuals consistently highlighted their emotional regulation efforts during these steps underscoring self-awareness as an important predictor of adaptive leadership behavior. Moreover, it is worth mentioning that some of the items of ALSA were also repeated in participants’ emotional regulation efforts especially Factor 1 and 3.

ALSE as another reflective mind component was validated and included in the model. Qualitative research did not provide any explanatory evidence about the role of ALSE in the reasoning process. However, its important role in the LMM10 suggests further investigation.

Qualitative Study Gives Insights about Surprises in the Quantitative Phase

Specific Individual Difference Matters. A notable surprise of this research pertained to the importance of specific individual variabilities in the baseline of leadership behavior within the adaptive leadership behavior model. As previously discussed, although the regression model indicated a significant relationship between ALB and reflective mind components, the correlation of residuals raised concerns about the model’s reliability and its ability to explain

certain patterns in the data. Therefore, LMM was employed, incorporating a random intercept alongside fixed effects and fixed intercept. This adjustment ultimately resulted in a robust model that predicts about 74% of variation in ALB. The unexpected nature of this finding came as a surprise for the researcher to question why including random intercepts mattered in predicting ALB. In other words, the inquiry shifted toward identifying specific individual difference in leadership behavior that is making a large difference among participants and should be accounted for in the model.

Delving into people's reasoning processes in the qualitative interviews offered a potential insight into this surprising finding. One possible explanation could be rooted in individuals' differences in their tolerance to dissonance, anxiety, and heat which is the first mental step described in the qualitative phase. This might also communicate the correlation between random intercept and MRM*ALSE interaction's random slope which gives further insights about what potential individual differences could be. Despite the theoretical discussion of dissonance within the framework and from various theoretical perspectives (Adaptive leadership, ELT, and DPT), it was not explicitly considered in the hypothesis. Hence, we assume that incorporating a subjective measurement in the model might lead to the convergence of the regression model and reducing the gap between R-squared conditions and R-squared marginal in LMM10.

What Are Leaders Adapting To? An Argument for the Importance of Relationality

Responding to this question serves as the culmination of this study. Throughout our research journey, which involved validating new scales, evaluating programs, and observing individuals' decision-making processes in exercising leadership, we gained valuable insights into Heifetz's adaptation. Heifetz's adaptation places adaptive challenge and the gap at the center, demands new strategies and abilities, occurs through experimentation, relies on diversifying the

risk, and generates loss (Heifetz, 1994; Heifetz et al., 2009). The reasoning process described in this research quantitatively and qualitatively illustrated these dimensions of adaptation. The ALB scale highlights these dimensions by introducing “surfacing conflict”, “seeking perspectives”, “creating space for multiple perspectives”, and “acting experimentally” as discerning factors of leadership behavior.

Adapting is burdensome. Our exploration of decision-making processes revealed a conflict-laden journey that requires individuals to adapt to negative emotions and cognitive burdens. In other words, both the quantitative and qualitative tracking of the reasoning process (From dissonance to assessment of failure) showed that exhibiting leadership behavior requires going beyond the comfort zone and adapting to uncomfortable emotional and cognitive state.

Can we bear this burden alone? The quantitative results through the LMM10 showed that how reflective mind components are necessary in exercising adaptive leadership particularly in resisting the status quo, regulating emotions, sustaining multiple interpretations, and putting them into actions. As discussed in the Chapter 2, there are two cognitive functions that are necessary in implementing the LMM10 model (Shifting from Type 1 to Type 2): Cognitive decoupling and working memory. Additionally, the qualitative phase highlighted the extent to which individuals should endure negative emotions and cognitive burdens throughout the process, emphasizing the necessity of emotional regulation. However, does everybody have these cognitive capacities/abilities anytime, anywhere? Especially, in the working environments with adaptive challenges like those were described in the interviews.

Humans are cognitively miserable, as asserted by Kahneman (2011) and Stanovich (2011), with our cognitive capacity constrained in managing these taxing functions. Algorithmic mind processes play a significant role in these functions, and not everyone may possess

sufficient cognitive capacity to fully engage these processes in all situations, particularly in adaptive leadership, which involves conflict management. This limitation is evident in the interviews, where it manifests as thinking errors such as SAC, override failure, and mindware gap. The LMM10 model suggests cognitive processes necessary for adaptive leadership behavior that may overwhelm individuals, making them impractical to run at all times and in all situations.

In response to this challenge, qualitative interviews have revealed the necessity of mobilizing others in adaptive leadership. Participants described a crucial mental step in their decision-making process, wherein they actively seek out others' perspectives. We have termed this step the "action space", where individuals, acting as leaders, create room for others to engage in discussing the adaptive challenge. This can be seen as a form of relational decoupling action. Instead of individuals solely conducting all simulations and interpretations with their limited cognitive capacity, they share the cognitive load by tapping into others' minds and cognitive capacities. This approach makes the transition to Type 2 processes less burdensome for individuals. Niemeyer (2019) referred to this approach to decision-making and reasoning as shared rationale, distributed reasoning, and intersubjective reasoning, while Niemeyer and Dryzek (2007) term it as intersubjective rationality. Thus, this research supports the importance of relationality by demonstrating the challenging process of exercising adaptive leadership individually.

Highlights of the Chapter

- The demographic characteristics of all sample datasets were outlined, with a majority of participants being women and individuals from underrepresented minority (URM) groups.
- Validation results for ALB, ALSA, ALSE, and ALSR were presented.
- Program evaluation findings for the YLE and LFC programs were provided. Generally, both programs showed positive effects on reflective mind components.
- An impact model for the YLE program was developed, revealing that KLC has a greater impact on individuals with higher rates of ALSR, ALSA, and ALSE.
- The LMM model illustrated the relationship between reflective mind components and ALB, confirming the replicability of DPT in the leadership development field.
- Participants' decision-making process was concretely described and mapped as a result of qualitative interviews. People reported involvement in four mental steps for exercising leadership.
- Connections between quantitative and qualitative results were made. Qualitative phase explained how and where in the decision-making process the reflective mind underlying cognitive mechanisms (ALSR, ALSA, and MRM) operate.

Chapter 5 - Conclusion & Recommendations

The goal of this research was to bridge the gap between learning and practicing leadership. Choosing KLC's 2023 path as the SRM, this study particularly focused on the three main gaps in the field of adaptive leadership development: *learning and practicing environment gap*, *measurement gap*, *teaching and assessment gap*. The longitudinal evaluation study, the engaged validation process of assessment tools, and the qualitative dive into people's leadership experiments were conducted specifically to address these three gaps. Moreover, this research journey helped to the development of assessment tools that are both valid and reliable and have practical applications for adaptive LDPs. These tools were not only developed but also utilized in a program evaluation study. Furthermore, their relevance was examined through an exploration of individuals' decision-making processes in their leadership interventions.

Achieving the objectives defined in Chapter 1, this study contributes to the existing literature in three key areas. First, it demonstrates the feasibility of measuring adaptive leadership competencies. Second, it elucidates a set of cognitive processes underlying the adaptive leadership behavior. Third, it offers tailored recommendations for LDPs to narrow the gap between learning and practicing leadership. This chapter delves into these three major contributions while revisiting the objectives, hypotheses, and question of the research. The results and findings of different phases of research will be reviewed. Finally, pertinent practical recommendations will be offered to adaptive leadership teachers/coaches and researchers.

Adaptive Leadership Measurement Tools

For several years, KLC's program evaluations lacked a research approach and was limited to measuring learning outcomes. Moreover, their most recent evaluation study (YLE 2021) was not theoretically aligned with their teaching framework. Now, with the development

of adaptation and leadership (or adaptive leadership) scales, the *teaching-measurement gap* discussed in the first chapter was addressed.

ALB Scale: Adaptive Leadership Behavior is Now Measurable

ALB unveiled crucial factors to be considered in defining adaptive leadership behavior. It identified the main elements of adaptive leadership that are not only theoretically sound but also backed by statistical evidence. Addressing the empirical ambiguity around the distinction between adaptive leadership intervention and other forms of relational leadership interventions, this research revealed four important discerning elements of adaptive leadership: seek perspectives, create space for multiple perspectives, surface conflict, and act experimentally.

These four elements fall under the overarching cycle of observation, interpretation, and intervention cycle. As outlined in Chapter 2, getting on the balcony to get a fresh perspective on the system's structures, culture, defaults, and power struggle is a pre-requirement in diagnosing the challenge. ALB items 7, 8, and 15 represent these observational actions on the balcony particularly identifying and learning about stakeholders and their power, influence, and loss. Regarding interpretation, ALB recognizes making multiple interpretations in its item 11 highlighting engaging others in considering multiple perspectives.

In terms of intervention, Chapter 2 listed three main behaviors: Regulating conflict, giving the work back, and holding steady. These three main actions are reflected in factors 2 and 3 of the ALB scale. Factor 2 highlights engaging and discussing conflicting priorities and competing commitments (Items 2, 3, 4, and 5) as well as raising the heat (Item 6), while Factor 3 underscores creating space for others' work. Regulating conflict is an idiosyncrasy of adaptive leadership which is reflected in the ALB scale. ALB indicates that engaging others in discussions about their competing priorities and conversations about the unspeakable constitutes a form of

leadership behavior. This can serve as a discerning factor in evaluating not only leadership behavior but also team-building capacities, as suggested by Avolio (1999).

Finally, Factor 4 of the ALB scale encapsulates the essence of adaptive leadership action, echoing Heifetz et al.'s (2009) assertion that “everything you do in leading adaptive change is an experiment” (p. 277).

ALB scale items show that contrary to the traditional understanding of leadership action as implementing a leader's vision, action in adaptive leadership framework manifests in two primary forms: mobilization and experimentation. Factors 1, 2, and 3 of the scale encompass behaviors through which individuals identify other stakeholders and invite them to participate in adaptive work. The essence of behavior in these relational actions relies on informed and conscious engagement.

Factor 4 also suggests that individuals engaged in adaptive work do not view their actions as final perfect solutions, but rather as experiments. Therefore, from the discerning perspective of ALB, a leader is someone who continually gathers feedback from engaged stakeholders and adjusts interventions based on experimentation and learning.

In summary, these factors portray adaptive work as an ongoing, process-oriented effort rather than a static, definitive solution to be implemented. This illuminates another aspect of adaptive leadership that sets it apart from other relational theories, especially transformational leadership. The mobilization nature of ALB items underscores its main purpose: progress on the challenge, as opposed to influence as the result of transformational leadership. While in transformational leadership, the outcome is influencing followers to work toward the leader's vision, ALB scale illustrates that in adaptive leadership the solution is not definitive and requires multiple perspectives (Factor 1) and experimentation (Factor 4) to be constructed. That is why

conflict is inevitable and managing conflict is a leadership behavior (Factor 2). In other words, despite transformational leadership, the object of work is not the follower that should be motivated to work towards the leader's goal, but the activity to mobilize others to make progress on the challenge and close the gap which is conflict-filled (Ospina, 2017).

Putting altogether, adaptive leadership behavior as represented in ALB items de-emphasize the influence-based relationship between individuals and removes the leader-follower transforming dynamics. The items reflect the inter-independent relationship between people rather than independent and dependent dynamics in transformational and transactional leadership (McCauley et al., 2006). ALB suggests that measuring leadership behavior should be focused on the activities that result in progress on the challenge not transforming, motivating, or stimulating individuals towards one's goal.

Adaptive Leadership Self-Awareness (ALSA): A Diagnostic Tool for Knowing Self

Knowing self as a system by understanding and identifying your triggers, loyalties, bandwidth, roles, purposes, and needs is an essential component of adaptive leadership (Heifetz, 1994; Heifetz & Linsky, 2002; Heifetz et al., 2009). Except for loyalties, ALSA captured all other aspects in three different factors: Awareness of defaults, awareness of roles and purposes, and awareness of needs. This shows that ALSA can be a reliable diagnostic tool for knowing self. Since awareness of emotions is key in emotional regulations according to Wrana (2014), responding to this survey at different stages of working on an adaptive challenge can help individuals to map their self-awareness and be able to navigate the conflict-filled process of adaptive leadership more consciously. For instance, this scale helps people assess their awareness of conflict tolerance, comfort zone, motivations, and their behavioral needs to work on a specific adaptive challenge in their organization or community.

ALSA yielded valuable insights in this research. It emerged as the second strongest predictor of ALB (both concurrently and longitudinally) showing that knowing self as a system is essential for intervening. The analysis also indicated that improving ALSA can impact ALB over time. Its convergence with ALSR as a goal orientation measurement tool suggests its role in translating multiple interpretations into actions. Additionally, ALSA's importance was frequently emphasized during the qualitative interviews, especially in regulating emotions and conflicts during the moments of dissonance and fostering engagement.

ALSR: Internalizing Locus of Control

Goal orientation dispositions play a pivotal role in driving leadership actions. According to Day et al. (2009), goal orientation is one of the five aspects of self-regulation, which is essential to leader development. This communicates the thinking dispositions that Stanovich (2011) argued are necessary for sustaining simulations and putting them into action. Informed by this, the researcher was prompted to consider the role of goal orientation in leadership intervention by posing the question "Is there any thinking disposition unique to adaptive leadership that serves as a goal orientation?"

ALSR includes some of the goal orientations driving any type of intervention. However, its second factor uniquely pertains to adaptive leadership. This factor shows that internalizing the locus of control-believing that they can lead rather than authority has the control of the work- by distinguishing leadership from authority is a driving mindset toward adaptive intervention. ALSR, as the strongest predictor of ALB after controlling for other variables and individual differences in the LMM model, provided empirical support for why separating leadership from authority is important and should be included in the model.

ALSR's unique role in predicting ALB is also evident in its discriminant validity with MRM. As confirmed by the qualitative interviews, this helps explain why some people may make multiple interpretations but fail to intervene. ALSR and MRM together demonstrate that even if someone has a high tendency to generate multiple interpretations (MRM), possessing an authority-dependent mindset (low ALSR) can hinder successful adaptive intervention. Conversely, in a contrasting scenario, having a self-authorizing mindset (High ALSR) without a tendency to make multiple simulations/interpretations (Low MRM) may not lead to a successful adaptive intervention. These findings underscore the intricate interplay between MRM and ALSR in generating effective adaptive leadership behaviors.

The researcher does not claim that ALB, ALSA, ALSE, and ALSR scales capture all dimensions of the adaptive leadership framework outlined in Heifetz's works. For instance, the ALSA scale does not capture knowing loyalties while ALB scale does not encompass managing self and thrive dimensions. Instead, this research highlights those aspects that were supported by valid, reliable, and generalizable statistical evidence. These scales underwent rigorous testing across different social groups (URM/non-URM, gender, Time, sector, education) to ensure their efficiency as measurement tools. Moreover, their application to assess the impact of KLC's path yielded significant results. It is believed that the elements identified in this research have enough discerning power to distinguish individuals with adaptive leadership capacity/mindset from those who lack.

Additionally, the assessment tools were designed in an engaged process. The scale items were designed in collaboration with the panel of experts and KLC's teaching team and staff who are going to use them to evaluate their programs. This collaboration of researchers, teaching team, and program funders and organizers would reduce the *teaching-measurement* gap.

A Model for Measuring Impact

This research took a comprehensive approach to address the gap in learning and practicing environments. Besides designing the assessment tool for measuring leadership behavior anchored in adaptive leadership concepts, the researcher endeavored to investigate the cognitive functions required in exercising this type of leadership. The researcher elaborated on these cognitive functions, particularly cognitive decouplings, both through testing quantitative hypotheses and responding to the qualitative question. In doing so, not only was the relationship between cognitive processes and adaptive leadership behavior revealed, but also a framework was developed that can showcase the unique impact model of different LDPs.

Reflective Components Predict Adaptive Leadership Behavior

The third hypothesis of this research focused on examining the relationship between adaptive leadership behavior and reflective mind components. Addressing this hypothesis was essential to fulfill the third objective of the research and provided recommendations for LDPs. Confirming this hypothesis necessitated a strong theoretical foundation supported by empirical data. This section summarizes how this objective was achieved.

This research delved deeply into the psychological underpinning of Heifetz's precondition for dealing with adaptive challenges: "will and skill to dance on the edge" (p. 26). By establishing theoretical connections between Heifetz's adaptive leadership and Stanovich's tripartite model of mind, this study emphasized the type of mind that involves cognitive processes underlying "will" in adaptive work.

Reflective mind processes as the translation for the folk psychology term of "Will" were focused on in this research to explore the mindsets that predict adaptive leadership behavior. Building this connection took theoretical excursions to the field of adult learning and

development and involved Kolb's (1974) experiential learning, Day et al.'s (2009) integrative leader development theory, and Kegan's CDT. These theoretical journeys culminated in the identification of common constructs that bridged Heifetz and Stanovich's frameworks, centering around the concept of adaptation; disturbance/inhibition, making multiple interpretations/simulations, and intervention/rational action to describe a reasoning process that leads to emergence of a new leadership behavior.

Empirical testing of this connection and process was the next step that required reliable measurement tools, some of which were not readily available. The constructs translated into five main variables (MRM, ALSE, ALSA, ALSR, and ALB) – to measure different steps of the reasoning process and together form a model to be tested in this research. The scales underwent a conservative and meticulous validation process, utilizing both quantitative (CV, EFA, CFA) and qualitative (Face validity, labeling constructs) methods. These scales were also tested on multiple sample populations to ensure generalizability.

Entering MRM, ALSA, ALSE, and ALSR as independent variables and ALB as a dependent variable into the regression model hypothesized in Hypothesis 3 did not converge into a practical model. Alternatively, an LMM model was created incorporating the IVs' fixed effects and random effects as well as DV's random and fixed intercepts. This model converged successfully explaining 74% of the variation in ALB. Controlling for specific individual differences in leadership behavior, the model confirmed the relationship between the reflective mind and adaptive leadership behavior showing that its components (IVs' fixed effects) and the interaction between MRM and ALSE together predict 35.5% of the variance in ALB. The interaction of ALSE and MRM and its random slope gave insights about the dynamics between algorithmic processes and reflective mind in predicting adaptive leadership behavior. Doing so,

LMM10 demonstrated that the reflective mind components deserve attention in LDPs to improve leadership behavior.

Moreover, convergence of the LMM10 with all its fixed and random effects showed that the replication of a cognitive psychology theory in the leadership field was successful.

Connecting these two fields of study, this research not only informed the LDPs with adaptive leadership teaching framework, but it also contributed to cognitive psychology and the debate on the value of dual process model of cognition. It particularly expands dual processing model to broader dynamic and changing environments where people deal with adaptive challenges as ill-defined problems (e.g., relational conflicts) that require learning, taking risks, and accepting change in priorities, beliefs, and habits (Eysenck & Keane, 2020; Heifetz et al., 2009).

KLC's Unique Model of Impact

In addressing Hypothesis 2, LMM9 was used as an assessment framework to examine KLC's model of impact. This research went beyond assessing whether KLC's path improves reflective mind variables and ALB – which it indeed does. The researcher was curious about how KLC path's impact on ALB varies across different levels of the reflective mind components (ALSA, ALSE, ALSR, and MRM).

Including Time as an attribute of the program in the model, it was observed that while KLC's impact on ALB is consistent across different levels of MRM, it is significantly dependent on the level of ALSA, ALSR, and ALSE. In other words, people with higher rates of ALSA, ALSR, and ALSE benefit more from the path. The implication of this finding will be discussed later in this chapter. Overall, LMM9 demonstrated its utility as a framework for modeling programs' impact mechanisms.

Mixed Methods: From Scanning to Zooming

While the quantitative phase played the role of a scanner, identifying some of the dynamics between reflective mind components and adaptive leadership behavior, qualitative interviews allowed us to zoom in on these dynamics. By delving deeper into individuals' concrete stories of leadership interventions and focusing on their emotions and thought processes, we gained more detailed insights into how these mechanisms operate. This helped us know more about leadership in practice and address the *training-transfer gap* by providing recommendations for the KLC teaching team.

The qualitative interviews illustrated how the scale items may resemble types of leadership behavior that people exercise, and the reflective mind components underlying them. Chapter 4 provided specific examples of leadership practices that speak to ALB, ALSA, ALSR, and MRM scale items. This way the qualitative phase not only corroborate the findings of scale validation process but also provided more insights about when and how the reflective mind components intervene in the reasoning process. For instance, the qualitative interviews highlighted the importance of ALSA by underlining its role at different stages of decision-making process. Furthermore, they elaborated on the differentiation between MRM and ALSR by delineating the reasoning processes involved in failed and successful interventions and identifying the specific thinking dispositions of each cognitive path. However, it is important to note that qualitative phase did not provide any evidence about the role of ALSE in the reasoning process toward adaptive leadership action despite the interesting finding in the quantitative phase about the MRM*ALSE interaction term and its random slope.

Additionally, the qualitative phase showed potential avenues for future research. For instance, in response to one of the surprises in the quantitative phase, it suggested “tolerance to

dissonance” as a potential explainer of the gap between R-squared marginal and conditional in LMMs. Moreover, the qualitative interviews unveiled a reflection/assessment phase that may initiate a new reasoning process to repair failed interventions, a phase that warrants further attention and exploration.

An Engaged Process of Theory Building

Different steps of this research illustrated Van de Ven’s (2009) diamond model of engaged scholarship. This model is comprised of four components three of which were practiced during this research.

- *Problem formulation:* This research was conducted based on KLC’s need in knowing what is working and what is not working in their programs. KLC, with more than 15000 alumni, lacked valid and reliable tools that are aligned with their theoretical framework.
- *Theory building:* This research took an abduction process in its multiple data collection and analyses phases. It started deductively by using Heifetz’s adaptive leadership framework to measure leadership behavior and Stanovich’s tripartite model of reasoning to track the reasoning process underlying this behavior. However, involving KLC’s teachers and participants, the interpretive phases of the measurements’ validation and qualitative interviews added an inductive dimension to this research which ultimately contributed to the initial theories. In sum, KLC’s team contributed to identifying the discerning factors of adaptive leadership behavior while participants outlined the reasoning process underlying adaptive leadership behavior.
- *Problem solving:* To respond to the problem formulated in this research, the teaching team at KLC received specific recommendations — which will be discussed next —

aimed at enhancing the effectiveness of their programs. These recommendations were provided during a focus group discussion that took place outside the timeline covered by this dissertation.

Recommendations

The qualitative and quantitative phases together illuminated the conflict-filled process of adaptive leadership intervention. From internal conflicts stemming from habit disruptions to relational conflicts while engaging others, there are several exit doors during this process that tempt individuals to abandon it. This research identified these exit doors using empirical qualitative and quantitative data and invites adaptive leadership teachers, coaches, and researchers to lock them by addressing the following recommendations.

Recommendations for Adaptive Leadership Teachers

Based on the findings that emerged at different phases of this research, four main recommendations are worth considering in teaching adaptive leadership:

- Based on the findings from LMM9 that indicate ALSA, ALSE, ALSR, and MRM together predict 35% of the leadership behavior, it is recommended that LDPs recognize the reflective mind components in their modules. In addition to building skills, emphasis should be placed on building a tendency to make multiple interpretations, a self-transforming mindset, and an awareness of defaults, roles, purposes, and needs. This becomes more important given that the predictive validity test showed influencing these tendencies during the program can have long-term effects. Improving ALSA may involve modules that my focus on mapping individuals' triggers during chaos and conflict, personal needs for change, specific behaviors need to be changed, formal and informal authority roles, and unmet personal needs that may impede intervention, besides

identifying the comfort zone boundaries. Enhanced tendency to make interpretations could be achieved through simulating potential loss as a result of single-interpretation-based interventions and risk assessment. Emphasizing authority vs leadership dynamics is highly recommended as data showed it can work as a driver mindset toward action.

- It is highly recommended that adaptive leadership teachers and coaches are familiar with the common decision-making process outlined here. This understanding can assist teachers in offering practical guidance for each mental step and its associated emotional conditions. This research has demonstrated that the emotion-laden and conflict-filled process of adaptive leadership can be a double-edged sword, potentially leading individuals toward either failure or success. This underscores the significance of leadership training, and the role educators play in the decision-making process. By informing individuals about what to expect at each stage and how to prepare for it, educators can better equip them for the challenges ahead. The following examples illustrate each step.

- Mental Step 1: This involves cognitive dissonance followed by negative emotionality experiencing anxiety, stress, shakiness, confusion, etc. Data suggests awareness of defaults, role, purpose, and needs can help participants navigate this step. Teachers/coaches can elaborate on these emotions and provide guidance on how self-awareness can help control these emotions.
- Mental Step 2: This is the analysis phase. We assume that building awareness about the taxonomy of thinking errors can be helpful for sustained adaptive action and reduce work avoidance tendency. Quantitative research showed that thinking dispositions for making multiple interpretations and intervening could be

different. Qualitative research provided more details about this phenomenon in practicing leadership. Interviews showed that the tendency to make multiple interpretations may not always lead to intervention due to failure in overriding habitual patterns. Moreover, making multiple interpretations or engaging others' perspectives may not necessarily result in ENGAGED intervention. People may make or engage multiple perspectives/interpretations but ultimately put their preferred interpretation into action — a phenomenon that is called SAC. Adaptive leadership teachers can make people more aware of these thinking errors while teaching leadership. Knowing this can increase the awareness of defaults particularly item ALSA 8 (*I know how I usually think and act*) and contributes to more successful adaptive interventions. Particularly for KLC's comfort zone exercise in YLE, this may justify delving deeper into thought barriers of expanding one's comfort zone. This research provided more details about what is happening in people's mind while moving outside their comfort zone.

- Mental Step 3: This step involves the decision to intervene. This research revealed important criteria for defining adaptive intervention. The four factors of the ALB scale and the decision-making process described in the qualitative phase show that an adaptive leadership intervention has an experimental and mobilization nature. Based on this idea, leadership is not implementing the solution individually, it is creating an engaging space where people collectively work toward the solution. As the ALB scale depicts, this phase also involves the skills to raise and manage conflict that — per qualitative results — requires emotional regulation. Besides highlighting the nature of adaptive leadership

action, teachers/coaches could emphasize the importance of self-awareness (defaults, role and purpose, and needs) in this phase. Elaborating on the relationship between ALSA and ALB in the quantitative phase, qualitative data suggests knowing self can help with navigating this typically conflict-filled process. Providing tools and techniques that help individuals reflect on their defaults, roles, purposes, and needs at the moments of dissonance (Mental Step 1) and engagement (Mental Step 3) might be helpful. It is worth noting that while knowing self (ALSA) is an important reflective mind component that leads to manage self, manage self itself was not emerged as a defining ALB factor. This means that “manage self” alone as a leadership behavior, advocated by KLC, does not distinguish a leadership actor from a non-actor. It is recommended that teachers/coaches communicate that managing self is only a step in leadership intervention process.

- Mental Step 4: This is the assessment/reflection phase which may happen after a non-adaptive intervention. These are the interventions mostly implemented by individuals themselves without engaging others. Some of these interventions are the result of SAC and some are the result of sticking to TASS. While these interventions might have negative consequences, participants reported possible repairing processes that can be initiated after this failure. Data shows that people may feel hesitant, scared, or unconfident to initiate this process. While dispositions are still involved as a commander to do the assessment and initiate potential repairing process, data indicates that awareness of needs emerges as a motivator when individuals recognize a pattern of past failed actions. Adaptive

leadership teachers are encouraged to acknowledge this reflection phase and provide helpful tools and guides for intervention assessment and repairing process. Moreover, building awareness of such a phase might help this reflection happen before taking action as a potential simulation within Mental Step 2.

- Besides their individual effect on ALB, the interaction between MRM and ALSE suggests that MRM can lead to more ALB at higher levels of ALSE. Since the evaluation study did not show increase in ALSE after the program and ALSE was not reflected in the qualitative interviews, it is recommended that KLC's teaching team pay special attention to this reflective mind component in their programs.
- KLC's model of impact for both LFC and YLE shows that individuals with more ALSA, ALSE, and ALSR are benefiting more from the path. It is strongly advised that KLC ensures a balanced and optimized impact on all participants. One potential action could involve enhancing participants' ALSA and ALSR within the "When Everyone Leads" program, which serves as the initial program of Path 2024.
- Evaluation study showed that the LFC program is more impactful on those who attended individually compared to group participants. While this may need further exploration by the teaching team, one potential interpretation provided in Chapter 4 highlights the importance of commitment to collective interventions. Without defining whose work it is to intervene, it is likely that participants leave the program without any commitment to group interventions and wait for others to intervene. It is recommended that the KLC teaching team encourages groups to design specific steps of interventions while defining roles and responsibilities.

Recommendations for Engaged Scholars

There are a few recommendations that engaged scholars might consider in developing instruments and conducting program evaluation studies while working with program providers and their participants:

- Formative program evaluation (not summative) needs to take an engaged process involving program providers and clients in order to have true implications for LDPs. It is likely that the evaluation studies that do not consider the needs of the program providers or involve them in the study fall short in informing and improving LDPs. In this research involving program providers in data collection contributed to the formative aspect of the evaluation.
- It is recommended that program evaluators establish sustained communication and interaction with program providers by forming a taskforce group/evaluation panel. The continuous interactions from the early stage of assessment tool development to the later stage of evaluation results interpretation could ensure the formative aspect of evaluation.
- Given the disparity between how impact assessment is perceived by scholars and non-academics (such as those in industry or companies), it is essential to engage in a collaborative process while developing assessment tools. This might ensure that the results of the evaluation are taken seriously. In doing so, it is recommended that the researchers' and program providers' commitments, loyalties, and goals being discussed beforehand. Researchers and LDP providers might have different goals of developing instruments that might affect the instrument design. In a tangible scenario, in item reduction phase of the scale validation, researchers may prioritize retaining scale items that establish connections between the instrument and existing field literature. In contrast,

program providers may focus solely on items that align closely with the content of their programs. This might even affect the instruments' naming. In another scenario, the researcher may prioritize the generalizability of the instruments across different contexts, whereas the program provider may be primarily focused on the effectiveness within their specific programs. These differences in goals should be addressed through discussions both before and during the research process.

- Building rapport with participants is crucial during data collection. It is important for participants to know that they are research collaborators and why the data (both quantitative and qualitative) are being collected. This rapport building can start from the program itself by highlighting the importance of data collection as part of development in the leadership journey. This can help with ensuring sufficient data, especially in the quantitative and longitudinal research.
- It is essential that neither the items in instruments nor the qualitative questions trigger the availability bias among participants, as this could undermine the reliability of evaluation results or place participants in a political predicament. For example, in this research, the research team removed the frequently repeated words in the LDPs from the scale items to mitigate bias. Furthermore, during the qualitative phase, the focus of questions shifted towards individuals' personal stories of leadership experiments, rather than emphasizing the program's contribution to their efforts. Asking direct questions about the program's impact might direct them toward a particular response and damage the trust.

Recommendations for Researchers

Improving Generalizability of the Scales

This research introduces ALB, ALSR, ALSE, and ALSA as validated measurement tools for evaluating any adaptive leadership programs. Although the generalizability of these scales has been tested across different social groups in Kansas, it is recommended that researchers evaluate these scales in cultural contexts other than Kansas and the US. Researchers in other countries may consider translating these scales into other languages and assessing their efficiency in assessing LDPs. Moreover, the adaptive leadership model (LMM10) and the regression model tested in this research could be evaluated for reliability in other contexts.

Moreover, due to item and factor structure change of scales particularly among Gen Z respondents, it is recommended the inconsistencies emerged in this specific population be investigated further.

Improving ALSR

Identified as a goal orientation scale and the strongest predictor of ALB, ALSR warrants further improvement. In developing ALSR, this research did not follow the best practices due to research limitations. However, it showed ALSR's strength and potential contributions to the adaptive leadership model and proposed potential ways to improve it. Knowing that ALSR can reflect on mastery and performance goal orientation and acknowledging that race can be a driver of intervention, including and testing these constructs as potential items of ALSR is highly recommended.

Exploring the Relationship Between Two Types of Minds

This research explored the effect of reflective mind components on adaptive leadership behavior. While there is room to further explore reflective components, our theoretical

framework suggests the relationship between the reflective mind and algorithmic mind is worth delving into. The LMM9 shows a 40% gap between R-squared marginal and conditional. We posit that this gap can potentially be addressed by adding algorithmic mind variables into the model. Moreover, our mixed-method analysis suggests that “tolerance to dissonance” can also be a potential explainer of this gap, especially given the loading of the MRM random slope in the model. That said, we propose three potential research studies related to the cognitive processes discussed in this research:

- Adding algorithmic mind components to the model identified in this research. This can involve measuring the effect of cognitive capacity and making simulation skills on ALB.
- Developing an instrument that measures tolerance to dissonance/heat in leadership and examining its contribution to LMM10.
- Investigating the negative correlation between ALB random intercept and $\text{MRM} \times \text{ALSE}$ random slope

References

- Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology*, 63(1), 1–18. <https://doi.org/10.1111/j.2044-8325.1990.tb00506.x>
- Auerswald, M., & Moshagen, M. (2019). How to determine the number of factors to retain in exploratory factor analysis: A comparison of extraction methods under realistic conditions. *Psychological Methods*, 24(4), 468–491. <https://doi.org/10.1037/met0000200>
- Avolio, B. J., Bass, B. M., & Jung, D. I. (1999a). Re-examining the components of transformational and transactional leadership using the Multifactor Leadership. *Journal of Occupational and Organizational Psychology*, 72(4), 441–462. <https://doi.org/10.1348/096317999166789>
- Avolio, B.J. (1999b). *Full leadership development*. Sage Publications.
- Avolio, B. J., & Gardner, W. L. (2005). Authentic leadership development: Getting to the root of positive forms of leadership. *The Leadership Quarterly*, 16(3), 315–338. <https://doi.org/10.1016/j.leaqua.2005.03.001>
- Avolio, B. J., & Hannah, S. T. (2008). Developmental readiness: Accelerating leader development. *Consulting Psychology Journal: Practice and Research*, 60(4), 331–347. <https://doi.org/10.1037/1065-9293.60.4.331>
- Avolio, B. J., Reichard, R. J., Hannah, S. T., Walumbwa, F. O., & Chan, A. (2009). A meta-analytic review of leadership impact research: Experimental and quasi-experimental studies. *The Leadership Quarterly*, 20(5), 764–784. <https://doi.org/10.1016/j.leaqua.2009.06.006>
- Avolio, B. J., Avey, J. B., & Quisenberry, D. (2010). Estimating return on leadership development investment. *The Leadership Quarterly*, 21(4), 633–644. <https://doi.org/10.1016/j.leaqua.2010.06.006>
- Bandura, A. (1986). *Social foundations of thought and action*. Prentice Hall.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. WH Freeman/Times Books/ Henry Holt & Co.
- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on psychological Science*, 1, 164–180.
- Bass, B.M. (1985). *Leadership and performance beyond expectations*. Free Press.
- Bass, B.M., & Avolio, B.J. (1990). Developing trans-formational leadership: 1992 and beyond. *Journal of European Industrial Training*, 14, 21–27. <http://dx.doi.org/10.1108/03090599010135122>

- Becker, H. (2001). Epistemology of qualitative research. In Emerson, R. (ed), *Contemporary field research: Perspectives and formulations* (pp. 53–71). Waveland Press.
- Bell, B. A., Morgan, G. B., Kromery, J. D., & Ferron, J. M. (2010). The impact of small cluster size on multilevel models: A Monte Carlo examination of two-level models with binary and continuous predictors. *JSM Proceedings, Survey Research Methods Section*, 1(1), 4057–4067.
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 1–62). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60024-6](https://doi.org/10.1016/S0065-2601(08)60024-6)
- Bhattacharya, K. (2017). *Fundamentals of qualitative research: A practical guide*. Routledge.
- Black, A. M., & Earnest, G. W. (2009). Measuring the Outcomes of Leadership Development Programs. *Journal of Leadership & Organizational Studies*, 16(2), 184–196. <https://doi.org/10.1177/1548051809339193>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, 1–18. <https://doi.org/10.3389/fpubh.2018.00149>
- Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation models: Present and future. A festschrift in honor of Karl Jöreskog* (pp. 139–168). Scientific Software International.
- Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes*, 97(2), 117–134. <https://doi.org/10.1016/j.obhdp.2005.03.002>
- Bryson, J.M & Patton, M.Q. (2015). Analyzing and engaging stakeholders. In K.E. Newcomer, K., H. P. Hatry, & J.S. Wholey (Eds.), *Handbook of practical program evaluation* (pp. 36–61). Jossey-Bass.
- Burns, J. M. G. (1978). *Leadership*. Harper & Row.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, 119(2), 197–253.
- Carless, S. A., Wearing, A.J., & Mann, L. (2000). A short measure of transformational leadership. *Journal of Business and Psychology*, 14(3), 389–405. <https://doi.org/10.1023/A:1022991115523>
- Carroll, B., Levy, L., & Richmond, D. (2008). Leadership as practice: Challenging the competency paradigm. *Leadership*, 4(4), 363–379. <https://doi.org/10.1177/1742715008095186>

- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10
- Chia, R., & Holt, R. (2006). Strategy as practical coping: A Heideggerian perspective, *Organization Studies*, 27(5), 635–55. <http://dx.doi.org/10.1177/0170840606064102>
- Chia, R., & MacKay, B. (2007). Post-processual challenges for the emerging strategy-as-practice perspective: Discovering strategy in the logic of practice. *Human Relations*, 60(1), 217–42. <https://doi.org/10.1177/0018726707075291>
- Chief Learning Officer (2017). *Follow the leaders(ship) spending*. <https://www.chieflearningofficer.com/2018/03/21/follow-the-leadership-spending/>
- Chrislip, D. D., & O'Malley, E. (2013). *For the common good: redefining civic leadership*. KLC Press.
- Collins, D. B., & Holton III, E. F. (2004). The effectiveness of managerial leadership development programs: A meta-analysis of studies from 1982 to 2001. *Human Resource Development Quarterly*, 15(2), 217–248. <https://doi.org/10.1002/hrdq.1099>
- Comrey A. L. (1988). Factor-analytic methods of scale development in personality and clinical psychology. *Journal of Consulting and Clinical Psychology*, 56(5), 754–761. <https://doi.org/10.1037//0022-006x.56.5.754>
- Crawford, J.A., Kelder, J.-A., (2019). Do we measure leadership effectively? Articulating and evaluating scale development psychometrics for best practice. *The Leadership Quarterly*, 30 (1), 133–144. <https://doi.org/10.1016/j.leaqua.2018.07.001>
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research* (pp. 209–240). Sage Publications.
- Cronbach, L.J., Nageswari, R., & Gleser, G.C. (1963). Theory of generalizability: A liberation of reliability theory. *The British Journal of Statistical Psychology*, 16, 137–163. <https://doi.org/10.1111/J.2044-8317.1963.TB00206.X>
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. Sage Publications.
- Drath, W. (2001). *The deep blue sea: Rethinking the source of leadership*. Jossey-Bass.
- Day, D. V. (2000). Leadership development: A review in context. *Leadership Development*, 11(4), 581–563. [https://doi.org/10.1016/S1048-9843\(00\)00061-8](https://doi.org/10.1016/S1048-9843(00)00061-8)
- Day, D. V., Fleenor, J. W., Atwater, L. E., Sturm, R. E., & McKee, R. A. (2014). Advances in leader and leadership development: A review of 25years of research and theory. *The Leadership Quarterly*, 25(1), 63–82. <https://doi.org/10.1016/j.leaqua.2013.11.004>

- 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(1), 133–156. <https://doi.org/10.1146/annurev-orgpsych-032414-111328>
- De Neys, W. (2018). *Dual process theory 2.0*. Routledge.
- Dewey, J. (1938). *Experience and education*. Macmillan.
- Denyer, D., & Turnbull James, K. (2016). Doing leadership-as-practice development. In J. A. Raelin (Ed.), *Leadership-as-Practice: Theory and application* (pp. 262-283). Routledge.
- Diener, Ed (1984). Subjective well-being. *Psychological Bulletin*, 95(3): 542–575. [doi:10.1037/0033-2909.95.3.542](https://doi.org/10.1037/0033-2909.95.3.542). PMID 6399758.
- Dinh, J. E., Lord, R. G., Gardner, W. L., Meuser, J. D., Liden, R. C., & Hu, J. (2014). Leadership theory and research in the new millennium: Current theoretical trends and changing perspectives. *The Leadership Quarterly*, 25(1), 36–62. <https://doi.org/10.1016/j.leaqua.2013.11.005>
- Dvir, T., Eden, D., Avolio, B. J., & Shamir, B. (2002). Impact of transformational leadership on follower development and performance: A field experiment. *Academy of Management Journal*, 45(4), 735–744. <https://www.jstor.org/stable/3069307>
- Emerson, Robert M., Rachel I. Fretz, and Linda L. Shaw. (2011). *Writing ethnographic fieldnotes*. University of Chicago Press.
- Evans JS. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annu Rev Psychol*, 59, 255–78. <https://doi.org/10.1146/annurev.psych.59.103006.093629>
- Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, 8(3), 223–241. <https://doi.org/10.1177/1745691612460685>
- Eysenck, M. W., & Keane, M. T. (2020). *Cognitive psychology: A student's handbook* (8th ed.). Psychology Press.
- Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175–191. <http://dx.doi.org/10.3758/BF03193146>
- Feehley, T., Plunkett, C. H., Bao, R., Choi Hong, S. M., Culleen, E., Belda-Ferre, P., Campbell, E., Aitoro, R., Nocerino, R., Paparo, L., Andrade, J., Antonopoulos, D. A., Berni Canani, R., & Nagler, C. R. (2019). Healthy infants harbor intestinal bacteria that protect against food allergy. *Nature Medicine*, 25(3), 448–453. <https://doi.org/10.1038/s41591-018-0324-z>
- Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

- Flora, D. B., & Flake, J. K. (2017). The purpose and practice of exploratory and confirmatory factor analysis in psychological research: Decisions for scale development and validation. *Canadian Journal of Behavioral Science / Revue Canadienne Des Sciences Du Comportement*, 49(2), 78–88. <https://doi.org/10.1037/cbs0000069>
- Gallucci, M. (2019). *GAMLj: General analyses for linear models*. [Jamovi module]. Retrieved from <https://gamlj.github.io/>.
- Giddens, A. (1979). *Central Problems in Social Theory: action, structure, and contradiction in social analysis*. Macmillan.
- Gorsuch, R. (1983). *Factor analysis (2nd ed.)*. Lawrence Erlbaum Associates.
- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological bulletin*, 103(2), 265–275. <https://doi.org/10.1037/0033-2909.103.2.265>
- Haber-Curran, P., & Tillapaugh, D. (2013). Leadership Learning through Student-Centered and Inquiry-Focused Approaches to Teaching Adaptive Leadership. *Journal of Leadership Education*, 12(1). [DOI:10.12806/V12/I1/R6](https://doi.org/10.12806/V12/I1/R6)
- Hannah, S. T., Avolio, B. J., Walumbwa, F. O., & Chan, A. (2012). Leader Self and Means Efficacy: A multi-component approach. *Organizational Behavior and Human Decision Processes*, 118(2), 143–161. <https://doi.org/10.1016/j.obhdp.2012.03.007>
- Haslett, J., & Hayes, K. (1998). Residuals for the linear model with general covariance structure. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 60, 201-215
- Hatcher, L. (1994). *Step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. SAS Institute.
- Heidegger, M. (1962). *Being and time*. Blackwell.
- Heifetz, R. A. (1994). *Leadership without easy answers*. Belknap Press of Harvard University Press.
- Heifetz, R. A., & Laurie, D. L. (1997). The work of leadership. *Harvard Business Review*, 75, 124–134.
- Heifetz, R., & Linsky, M. (2002). *Leadership on the line*. Harvard Business Review.
- Heifetz, R. A., Linsky, M., & Grashow, A. (2009). *The practice of adaptive leadership*. Harvard Business Review Press.

- Hiller, N. J. (2005). *An examination of leadership beliefs and leadership self-identity: Constructs, correlates, and outcomes* [Doctoral dissertation, Pennsylvania State University].
- Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. *Journal of Management*, 21(5), 967–988. <https://doi.org/10.1177/014920639502100509>
- Hogarth, R. M., Lejarraaga, T., & Soyer, E. (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science*, 24(5), 379–385. <https://doi.org/10.1177/0963721415591878>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, 89(5), 755–768. <https://doi.org/10.1037/0021-9010.89.5.755>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Katz, J. (2001). A theory of qualitative methodology: The social system of analytic fieldwork. In Robert, E. (Ed), *Contemporary field research: A collection of readings* (pp. 127–148). Waveland Press Inc.
- Kegan, R. (1982). *The evolving self: Problem and process in human development*. Harvard University Press.
- Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Harvard University Press.
- Kegan, R., & Lahey, L. L. (2009). *Immunity to change*. Harvard Business Review Press.
- Kelman, H. C. (1958). Compliance, identification, and internalization: Three Processes of attitude change. *Journal of Conflict Resolution*, 2, 51-60. <https://doi.org/10.1177/002200275800200106>
- Kim, T., Sydnes, A. K., & Batalden, B.-M. (2021). Development and validation of a safety leadership Self-Efficacy Scale (SLSES) in maritime context. *Safety Science*, 134, 105031. <https://doi.org/10.1016/j.ssci.2020.105031>
- Kirton, M. J. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology*, 61(5), 622–629. <https://doi.org/10.1037/0021-9010.61.5.622>
- Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization: Theory and research* (pp.347–480). Rand McNally.

- Kolb, D. A., & Fry, R. (1975). Towards an applied theory of experiential learning. In C. Cooper (Ed.), *Theories of Group Process*. John Wiley.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.
- Kolb, D. A. (2014). *Experiential learning: Experience as the source of learning and development*. FT Press.
- Kyriazos, T. A. (2018). Applied psychometrics: Sample size and sample power considerations in factor analysis (EFA, CFA) and SEM in general. *Psychology*, 9(8), 2207–2230. <https://doi.org/10.4236/psych.2018.98126>
- Lacerenza, C. N., Reyes, D. L., Marlow, S. L., Joseph, D. L., & Salas, E. (2017). Leadership training design, delivery, and implementation: A meta-analysis. *Journal of Applied Psychology*, 102(12), 1686–1718. <https://doi.org/10.1037/apl0000241>
- Leadership IQ (2020). *2020 State of leadership development*. <https://www.leadershipiq.com/blogs/leadershipiq/leadership-development-state>
- Leslie, A. M. (1987). Pretense and representation: The origins of “Theory of Mind.” *Psychological Review*, 94(4), pp. 412–426. <https://doi.org/10.1037/0033-295X.94.4.412>
- Liden, R. C., Wayne, S. J., Zhao, H., & Henderson, D. (2008). Servant leadership: Development of a multidimensional measure and multi-level assessment. *The Leadership Quarterly*, 19(2), 161–177. <https://doi.org/10.1016/j.leaqua.2008.01.006>
- Lim, S., & Jahng, S. (2019). Determining the number of factors using parallel analysis and its recent variants. *Psychological Methods*, 24(4), 452–467. <https://doi.org/10.1037/met0000230>
- Locke, E. A. (2007). The case for inductive theory building. *Journal of Management*, 33(6), 867–890. <https://doi.org/10.1177/0149206307307636>
- Loew, L., & Wentworth, D. (2013). *Leadership: The state of development programs: Research based industry perspective*. Brandon Hall Group.
- Lord, R. G., & Hall, R. J. (2005). Identity, deep structure and the development of leadership skill. *The Leadership Quarterly*, 16(4), 591–615. <https://doi.org/10.1016/j.leaqua.2005.06.003>
- Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods*, 49(4), 1494–1502. <https://doi.org/10.3758/s13428-016-0809-y>
- Lynn M. R. (1986). Determination and quantification of content validity. *Nursing research*, 35(6), 382–385.
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling.

- Methodology*, 1(3), 86-92. [DOI: 10.1027/1614-1881.1.3.86](https://doi.org/10.1027/1614-1881.1.3.86)
- Marcus, G. E. (2002). *The sentimental citizen: Emotion in democratic politics*. Pennsylvania State University Press.
- Marion, R., & Uhl-Bien, M. (2001). Leadership in complex organizations. *The Leadership Quarterly*, 12(4), 389–418. [https://doi.org/10.1016/S1048-9843\(01\)00092-3](https://doi.org/10.1016/S1048-9843(01)00092-3)
- Martin, R.J. (1992). Leverage, influence and residuals in regression models when observations are correlated. *Communications in Statistics-Theory and Methods*, 21, 1183–1212. <https://doi.org/10.1080/03610929208830840>
- Marsh, H. W., & Hau, K. T. (1999). Confirmatory factor analysis: strategies for small sample sizes. *Statistical Strategies for Small Sample Research*, 1, 251–284.
- McCauley, C. D., Drath, W. H., Palus, C. J., O'Connor, P. M. G., & Baker, B. A. (2006). The use of constructive-developmental theory to advance the understanding of leadership. *The Leadership Quarterly*, 17(6), 634–653. <https://doi.org/10.1016/j.leaqua.2006.10.006>
- McGuire, J. B., & Rhodes, G. B. (2009). *Transforming your leadership culture*. Jossey-Bass.
- Merton, R. (1987). Three fragments from a sociologist's notebooks. *Annual Review of Sociology*, 13, 1-28.
- Meteyard, L., & Davies, R. A. I. (2020). Best practice guidance for linear mixed-effects models in psychological science. *Journal of Memory and Language*, 112, 104092. <https://doi.org/10.1016/j.jml.2020.104092>
- Meyers, R. (1999). The beginnings of pragmatism: Peirce, Wright, James, Royce. In R. Popkin (Ed.), *The Columbia history of Western philosophy* (pp. 592–600). Columbia University Press.
- Miscenko, D., Guenter, H., & Day, D. V. (2017). Am I a leader? Examining leader identity development over time. *The Leadership Quarterly*, 28(5), 605–620. <https://doi.org/10.1016/j.leaqua.2017.01.004>
- Murphy, S. E. (1992). *The contribution of leadership experience and self-efficacy to group performance under evaluation apprehension* [Doctoral dissertation, University of Washington].
- Murphy, J. I., Weaver, N. E., & Hendricks, A. E. (2022). Accessible analysis of longitudinal data with linear mixed effects models. *Disease Models & Mechanisms*, 15(5), 1–8. <https://doi.org/10.1242/dmm.048025>
- Natanson, M. (1973). *Edmund Husserl: Philosopher of infinite tasks*. Northwestern University Press.

- Neider, L. L., & Schriesheim, C. A. (2011). The authentic leadership inventory (ALI): Development and empirical tests. *The Leadership Quarterly*, 22(6), 1146–1164. <https://doi.org/10.1016/j.leaqua.2011.09.008>
- Newcomer, K. E., Hatry, H. P., & Wholey, J. S. (2015). *Handbook of practical program evaluation* (4th ed.). Jossey-Bass.
- Ng, J. C. (2021). Refuse thy name? Critical reflections on the convention of masking place in rural educational research. *International Journal of Qualitative Studies in Education*, 34(9), 790–799. <https://doi.org/10.1080/09518398.2021.1956624>
- Niemeyer, S., and Dryzek, J. (2007). The ends of deliberation: Metaconsensus and intersubjective rationality as deliberative ideals. *Swiss Political Science Review*, 13(4), 497–526. <https://doi.org/10.1002/j.1662-6370.2007.tb00087.x>
- Niemeyer, S. (2019). *Intersubjective reasoning in political deliberation*. Center for Deliberative Democracy & Governance.
- Northouse, P. G. (2016). *Leadership, theory and practice*. Sage Publications.
- Nunnally, J.C., & Bernstein, I.H. (1994). The assessment of reliability. *Psychometric Theory*, 3, 248-292.
- O'Leonard, K., & Krider, J. (2014). *Leadership development factbook 2014: Benchmarks and trends in U.S. leadership development*. BERSIN by Deloitte.
- O'Malley, E., & Cebula, A. (2015). *Your leadership edge*. KLC Press.
- O'Malley, E., & McBride, J.F (2023). *When everyone leads*. KLC Press.
- Osborne, W.J. (2019). What is rotating in exploratory factor analysis? *Practical Assessment, Research, and Evaluation*, 20(2), 1–7. <https://doi.org/10.7275/hb2g-m060>
- Ospina, S. M., & Uhl-Bien, M. (2012). Exploring competing bases for legitimacy in contemporary leadership studies. In Uhl-Bien, M., and Ospina, S. M. (Eds.), *Advancing relational leadership: A dialogue among perspectives* (pp. 1-40). Information Age Publishing.
- Ospina, S. M. (2016). Collective Leadership and Context in Public Administration: Bridging Public Leadership Research and Leadership Studies. *Public Administration Review*, 77(2), 275–287. <https://doi.org/10.1111/puar.12706>
- Pellerin, N., & Raufaste, E. (2020). Psychological Resources Protect Well-Being During the COVID-19 Pandemic: A Longitudinal Study During the French Lockdown. *Frontiers in psychology*, 11, 590276. <https://doi.org/10.3389/fpsyg.2020.590276>
- Petty, R.E., & Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, 19, 123-205. <https://doi.org/10.1016/S0065->

- Piaget, J. (1954). *The construction of reality in the child*. Basic Books.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Pugh, A. J. (2013). What good are interviews for thinking about culture? Demystifying interpretive analysis. *American Journal of Cultural Sociology*, 1(1), 42–68. <https://doi.org/10.1057/ajcs.2012.4>
- Raei, M. (2018). *Development and validation of the adaptive leadership with authority scale* [Doctoral dissertation, Antioch University]. <https://aura.antioch.edu/etds/446>
- Raelin, J. (2011). From leadership-as-practice to leaderful practice. *Leadership*, 7(2), 195–211. <https://doi.org/10.1177/1742715010394808>
- Ratner, B. (2009). The correlation coefficient: Its values range between +1/–1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing*, 17(2), 139–142. <https://doi.org/10.1057/jt.2009.5>
- R Core Team (2022). *R: A Language and environment for statistical computing*. (Version 4.1) [Computer software]. Retrieved from <https://cran.r-project.org>. (R packages retrieved from CRAN snapshot 2023-04-07).
- Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning, *Policy Sciences*, 4, 155–169. <https://doi.org/10.1007/BF01405730>
- Rosenkranz, G.K. (2015). Analysis of cross-over studies with missing data. *Statistical Methods in Medical Research*, 24, 420–433.
- Rumi, J. (13th century). In it, what's in it [*Fihi ma fihi*].
- Russell, B. (2001). *The problems of philosophy*. Oxford University Press.
- Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed). Sage Publications.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, 84, 1–66. <https://doi.org/10.1037/0033-295X.84.1.1>

- Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 118–137). Sage Publications.
- Schwandt, T.A. (2007). *The Sage dictionary of qualitative inquiry* (3rd ed.). Sage Publications.
- Seibel, M., Kaufman, E. K., Cletzer, D. A., & Elliott-Engel, J. (2023). Advancing adaptive leadership through ADAPTION-INNOVATION theory: Enhancements to the holding environment. *Journal of Leadership Studies*, 17(1), 23–29.
<https://doi.org/10.1002/jls.21841>
- Shams, K., Steffensmeier T., Dailey, S. (In revision). Unveiling the impact: A mixed-method inquiry into the impact of leadership development programs. *Journal of Leadership Education*, Vol (Issue), pp.
- Small, M. (2009). How many cases do I need? on science and the logic of case selection in field-based research. *Ethnography*, 10(1), 5–38. <https://doi.org/10.1177/1466138108099586>
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Lawrence Erlbaum Associates Publishers.
- Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning*, 13(3), 225–247. <https://doi.org/10.1080/13546780600780796>
- Stanovich, K. E. (2008). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 55–88). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199230167.003.0003>
- Stanovich, K. E. (2011). *Rationality and the reflective mind*. Oxford University Press.
- Stanovich, K. E., & West, R.F. (2011). Taxonomy of rational thinking problems. In K. E. Stanovich (Ed.), *Rationality and the reflective mind* (pp. 95–119). Oxford University Press.
- Stanovich, K. E., West, R.F., & Toplak, M. E. (2011). The assessment of rational thought. In K. E. Stanovich (Ed.), *Rationality and the reflective mind* (pp. 191–246). Oxford University Press.
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.
- Tavory, I., & Timmermans, ST. (2014). *Abductive analysis; Theorizing qualitative research*. The University of Chicago Press.
- Taylor, S.N. (2010). Redefining leader self-awareness by integrating the second component of self-awareness. *Journal of Leadership Studies*, 3, 57–68. <https://doi.org/10.1002/jls.20139>

- The Jamovi project (2023). *Jamovi*. (Version 2.4) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- Thoresen, M., & Laake, P. (2007). A simulation study of statistical tests in logistic measurement error models. *Journal of Statistical Computation and Simulation*, 77(8), 683–694. <https://doi.org/10.1080/10629360600589541>
- Timmermans, St., & Tavory, I. (2012). Theory construction in qualitative research: From grounded theory to abductive analysis. *Sociological Theory*, 30(2), 167–186. <https://doi.org/10.1177/0735275112457914>
- Tinbergen, N. (1963). On aims and methods in ethology. *Zeitschrift für Tierpsychologie*, 20, 410–433. <https://doi.org/10.1111/j.1439-0310.1963.tb01161.x>
- Tolich, M. (2004). Internal confidentiality: When confidentiality assurances fail relational informants. *Qualitative Sociology*, 27(1), 101–106.
- Torbert, B., & Associates. (2004). *Action inquiry: The secret of timely and transforming leadership*. Berrett-Kohler
- Tracy, S. J. (2010). Qualitative quality: Eight “Big-Tent” criteria for excellent qualitative Research. *Qualitative Inquiry*, 16(10), 837–851. <https://doi.org/10.1177/1077800410383121>
- Training Industry, Inc. (2022). *State of leadership training market*. <https://store.trainingindustry.com/products/the-state-of-the-leadership-training-market>
- Torbert, B., & Associates. (2004). *Action inquiry: The secret of timely and transforming leadership*. Berrett-Kohler.
- Uhl-Bien, M., Marion, R., and McKelvey, B. (2007). Complexity leadership theory: Shifting leadership from the industrial age to the knowledge era. *The Leadership Quarterly*, 18 (4), 298–318. <https://doi.org/10.1016/j.leaqua.2007.04.002>
- Uhl-Bien, M., & Marion, R. (2009). Complexity leadership in bureaucratic forms of organizing: A meso model. *The Leadership Quarterly*, 20(4), 631–650. <https://doi.org/10.1016/j.leaqua.2009.04.007>
- Uhl-Bien, M. and Arena, M. (2017). Complexity leadership: Enabling people and organizations for adaptability. *Organizational Dynamics*, 46, 9–20. <https://doi.org/10.1016/j.orgdyn.2016.12.001>
- Van de Ven, A. H. (2007). *Engaged scholarship: A guide for organizational and social research*. Oxford University Press.
- Van Manen, M. (2014). *Phenomenology of practice: Meaning-giving methods in phenomenological research and writing*. Left Coast Press.

- Velleman, J. D. (1992). What happens when someone acts? *Mind*, 101(403), 461–481.
<http://www.jstor.org/stable/2253898>
- Vogel, R. (2012). The visible colleges of management and organization studies: A bibliometric analysis of academic journals. *Organization Studies*, 33(8), 1015–1043.
<https://doi.org/10.1177/0170840612448028>
- Vogel, B., Reichard, R. J., Batistič, S., & Černe, M. (2021). A bibliometric review of the leadership development field: How we got here, where we are, and where we are headed. *The Leadership Quarterly*, 32(5), 101381. <https://doi.org/10.1016/j.leaqua.2020.101381>
- Walumbwa, F. O., Avolio, B. J., Gardner, W. L., Wernsing, T. S., & Peterson, S. J. (2008). Authentic leadership: Development and validation of a theory-based measure. *Journal of Management*, 34(1), 89–126. <https://doi.org/10.1177/0149206307308913>
- Williams, D. (2005). *Real leadership: Helping people and organizations face their toughest challenges*. Berrett-Koehler.
- Weiss, R. (1995). *Learning from strangers: The art and method of qualitative interview studies*. American Sociological Association.
- Weng, J., & Hubbard, L. (2022). Understanding the impact of case-in-point courses in graduate leadership programs. *The Journal of Leadership Education*, 21(2), 51–73.
<https://doi.org/10.12806/V21/I2/R4>
- Wolfinger, R., Tobias, R., & Sall, J. (1991). *Mixed models: A future direction*. 16th SAS Users Group Conference.
- Yukl, G., & Mahsud, R. (2010). Why flexible and adaptive leadership is essential. *Consulting Psychology Journal: Practice and Research*, 62(2), 81–93.
<https://doi.org/10.1037/a0019835>

Appendix A - Item Generation

Appendix A.1. ALB Items Generated from the Literature

Diagnosing Challenges

1. I engage myself in conversation about why I have not done enough to make change happen. (Adaptive challenge archetype 1)
2. I engage people in conversation about why the organization and themselves have not done enough to make change happen. (Adaptive challenge archetype 1)
3. I discuss my priorities with others. (Adaptive challenge archetype 2)
4. I engage myself to speak the unspeakable in the organization's/community's conversations. (Adaptive challenge archetype 3)
5. I engage myself in conversations in which my competing commitments are discussed. (Adaptive challenge archetype 2)
6. I engage people to speak the unspeakable in the organization's/community's conversations. (Adaptive challenge archetype 3)
7. I notice avoidance and give heads up about the avoidance tactics. (Adaptive challenge archetype 4)

Diagnosing Political Landscape

8. I identify my organization's/community's different stakeholders and gather information about their desired outcomes, levels of engagement, and power or influence.
9. I identify my organization's/community's different stakeholders and gather information about their values, loyalties, losses.

Managing Self

10. I present a fixed version of myself in different conditions and times based on my values and loyalties. (R)
11. I identify my triggers when I get overwhelmed/angry.
12. I tolerate and generate uncertainty and conflict.
13. I share my vulnerabilities at work with my confidants (e.g., consultant, family members, close friends)

Purposes

14. I try different ways to keep my purposes alive for acting leadership.
15. My purposes of acting leadership are not negotiable. (R)

Energizing Others

16. I help people move from simple to complex response.
17. I engage myself in making multiple tough interpretations.
18. I engage people in making multiple tough interpretations.

Intervening Skillfully: Design effective interventions

- 19. I assess the situation before heating up an issue.
- 20. I consult my intervention with others before taking action.
- 21. I let people in the organization/community work with my idea and give me feedback.
- 22. I analyze different stakeholders related to the issue in my organization/ community.
- 23. I help people with their resistance to change by providing more information about the problem and their constituencies' loss and fears.

Acting Politically

- 24. I strengthen my relationships.
- 25. I score some early wins.
- 26. I address interests unconnected to different difficult challenges.
- 27. I take small steps of my intervention ideas.
- 28. I get allies for addressing the issue in the organization/community.
- 29. I prepare my boss for my intervention in difficult organizational/community challenges.
- 30. I take responsibility of my interventions' casualties.
- 31. I engage the voices of naysayers and opponents of my idea.

Orchestrating Conflict

- 32. I regulate the heat in the working environment (increase or lower the heat). (Holding environment)
- 33. I select participants successfully to make progress on challenges. (Holding environment)
- 34. I usually start working on challenges and continue to lead alone. (R) (Give work back to people)

Engage Courageously

- 35. I do not risk myself being incompetent. (R)
- 36. I strengthen my capacity to make tougher decisions.
- 37. I prepare my organization for possible failure in my actions.
- 38. I conduct small experiments to test my ideas.

Engaging Courageously: Inspiring others (listening and speaking from heart)

- 39. I listen with more judgement than curiosity. (R)
- 40. I allow more for silence in meetings, arguments, and discussions.
- 41. I speak while holding my emotions.

Running Experiments

- 42. I run and test experimental interventions in my organization/community to make progress on difficult challenges.
- 43. I go beyond what is authorized to me formally in my organization/community.
- 44. I turn up the heat on an issue in my organization/community without making myself the issue.

Thrive

45. I create mental and physical sanctuaries to recover myself from work. (Taking care of self)
46. I usually look for sources of warmth, recognition, and praise in my organization/community.
47. I have a balanced portfolio for my work and life.

Appendix A.2. ALSE Items Generated from the Literature

I can...

1. Diagnose situations containing wicked problems.
2. Diagnose how my organization/community is responding to difficult challenges.
3. Diagnose the context underlying challenges.
4. Manage myself in conflict and uncertainty.
5. Energize others in making progress on difficult challenges.
6. Design effective interventions to make progress on difficult challenges.
7. Act politically in dealing with challenges.
8. Manage conflicts.
9. Engage courageously to make progress on challenges.
10. Inspire others.
11. Run experiments to make progress on challenges.

Appendix A.3. ALSA Items Generated from the Literature

1. I know what complex forces are influencing my behavior and choices.
2. I know what role I am playing in the larger organization/community.
3. I know what purposes I am seeking to serve.
4. I know what changes I need to make in myself to lead more effectively.

Loyalties

5. I am aware of my feelings of obligation toward my colleagues that may come into conflict while dealing with difficult challenges.
6. I am aware of my loyalties (people, groups, beliefs, places) that can support/impede my progress on difficult challenges.

Personal Tuning

7. I am aware of my unmet personal needs.
8. I am aware of my susceptibility to carrying other people's hopes and expectations
9. I am aware of my level of tolerance for the chaos, conflict, and confusion that accompany challenges.
10. I am aware of my triggers when I am overwhelmed/angry.

Bandwidth

11. I am aware of my mental comfort zone.

My Roles

12. I am aware of the distinction between my role (e.g., colleague, spouse, friend, sibling) and myself.
13. I am aware of what formal authorizers authorize me within the organization/community. (Formal authority)
14. I am aware of my role when people look to me in some ways to meet their needs. (informal authority)
15. I am aware of whose support I may need to accomplish my job. (informal authority)

Purpose

16. I am aware of my specific life purposes.
17. I can see realities of the world as truths not assumptions. ®
18. I have strong purpose(s) to practice leadership.
19. I do not negotiate my purposes with others. ®
20. I am aware of my aspirations for acting leadership.
21. I am aware of my ambitions for acting leadership.

Engage Courageously

22. I give permission to myself to fail in my actions.

Run Experiments

23. I usually fall in love with my ideas. ®
24. I unusually am aware of my part of the mess in my organization's/community's challenge.
25. I am aware of my own incompetence.

Appendix A.4. ALSR Items Generated from the Literature

1. Bringing change to my community/work is my personal priority.
2. Bringing change is my assigned responsibility in my work/community.
3. I practice leadership because a person should always consider new ways of solving a problem.
4. I believe we should look to our authorities for solving problems. ®
5. Difficulties can usually be overcome through waiting for good fortune, rather than thinking about the problem. ®
6. Even if my work/community environment had been different, I probably would have the same tendency for change.
7. I believe things must be the way they are. ®

Appendix B - Final List of Items After Focus Group 1

Appendix B.1. ALB Final List After Focus Group 1

- ALB1. I engage others in conversation about why they have not done enough to make change happen.
- ALB2. I discuss my conflicting priorities with others.
- ALB3. I engage others in conversations about our organization's/community's conflicting priorities.
- ALB4. I engage myself in conversations in which my competing commitments are discussed.
- ALB5. I engage others to say what is usually unspeakable in the organization/community.
- ALB6. I call attention when my organization/community resists confronting challenges.
- ALB7. I identify stakeholders and gather information about their power or influence.
- ALB8. I identify stakeholders and gather information about what they are afraid of giving up.
- ALB9. I tolerate uncertainty and conflict.
- ALB10. I consider multiple perspectives about problems at work/within the community.
- ALB11. I engage others in considering multiple perspectives about problems at work/within the community.
- ALB12. I assess the situation before creating tension about an issue.
- ALB13. I consult with others before taking action.
- ALB14. I let people in the organization/community work with my idea and give me feedback.
- ALB15. I try to learn about different stakeholders related to the issue in my organization/community.
- ALB16. I help people overcome resistance to change by providing information and acknowledging negative consequences.
- ALB17. I implement my intervention ideas with small steps.
- ALB18. I limit risk by letting my boss know how I intend to intervene.
- ALB19. I take responsibility for any damage resulting from my interventions.
- ALB20. I engage the voices of naysayers and opponents.
- ALB21. I regulate the intensity in the working environment.
- ALB22. I build my capacity to make tougher decisions.
- ALB23. I try new things to test my ideas.
- ALB24. I allow for silence in meetings, arguments, and discussions.
- ALB25. I speak clearly even while managing my emotions.
- ALB26. I regularly run tests to make progress on difficult challenges.
- ALB27. I prepare my organization for possible failure in my actions.
- ALB28. I regularly go beyond what I am formally authorized to do in my organization/community.
- ALB29. I generate tensions on important issues in my organization/community without making myself the issue.
- ALB30. I take time to rest and renew, physically and mentally.

Appendix B.2. ALSE Final List After Focus Group 1

- ALSE1. I can analyze situations containing complex challenges.
- ALSE2. I can analyze how my organization/community is responding to difficult challenges.
- ALSE3. I can analyze the root-causes of a challenge.
- ALSE4. I can manage myself amidst conflict and uncertainty.
- ALSE5. I can energize others to make progress on difficult challenges.
- ALSE6. Design effective interventions to make progress on difficult challenges.
- ALSE7. I am confident of my ability to manage conflict.
- ALSE8. Engage bravely to make progress on challenges.
- ALSE9. I feel that I know how to inspire others.
- ALSE10. Test ideas to make progress on challenges.

Appendix B.3. ALSA Final List After Focus Group 1

- ALSA1. I know what role I play in relation to an important challenge in my organization/community.
- ALSA2. I regularly assess my motivations for trying to create change.
- ALSA3. I know which of my behaviors I need to change to lead more effectively.
- ALSA4. I am aware of my loyalties (people, groups, beliefs, places) that could impede my progress on complex challenges.
- ALSA5. I am aware of my unmet personal needs that could impede my exercise of leadership.
- ALSA6. I am aware of my level of tolerance during moments of chaos, conflict, and confusion in my organization/community.
- ALSA7. I am aware of the triggers that make me angry/overwhelmed.
- ALSA8. I am aware of the way I usually think and act.
- ALSA9. I am aware of the distinction between my role (e.g., colleague, partner, friend, sibling) and who I am.
- ALSA10. I am aware of what I am formally authorized to do within the organization/community.
- ALSA11. I have strong reason(s) to practice leadership.
- ALSA12. I give myself permission to fail.
- ALSA13. I tend to fall in love with my ideas. ®
- ALSA14. I realize that I myself may contribute in some ways to the challenge in my organization/community.

Appendix B.4. ALSR Final List After Focus Group 1

- ALSR1. Creating change at work/in my community is my personal priority. (*Mastery goal orientation*)
- ALSR2. Creating change is part of my job description. (*Performance goal orientation*)
- ALSR3. I practice leadership because a person should always consider new ways of solving a problem. (*Mastery goal orientation*)
- ALSR4. I believe we should look people in positions of authority to solve problems. ® (*Mastery goal*)

- ALSR5. Difficulties can usually be overcome by waiting for good fortune, rather than thinking about the problem. ® (*Mastery goal orientation*)
- ALSR6. Even if my work/community environment had been different, I probably would have had the same tendency for change. (*Mastery goal orientation*)
- ALSR7. I believe there is little I can do to change the way things are. ® (*Performance goal orientation*)

Appendix C - Content Validity Results

Due to the large size of the tables, they could not be included here. You can find the content validity tables at this link: <https://shorturl.at/dp467>

Appendix D - EFA Results for ALB Scale

Appendix D Table D.1. Factor Loadings

	Factor						Uniqueness
	1	2	3	4	5	6	
LB3	0.845						0.394
LB2	0.822						0.486
LB4	0.598						0.437
LB5	0.491						0.551
LB6	0.479						0.482
LB1	0.466		0.357				0.637
LB7		1.011					0.306
LB8		0.923					0.345
LB15		0.866					0.444
LB20							0.699
LB9							0.925
LB26			0.992				0.320
LB27			0.543			0.395	0.585
LB17			0.371				0.697
LB22			0.329				0.537
LB25			0.303				0.674
LB10				1.014			0.339
LB11				0.817			0.449
LB23			0.303	0.329			0.600
LB18							0.777
LB19							0.786
LB13					0.598		0.785
LB14					0.563		0.515
LB24					0.563		0.771
LB16					0.509		0.613
LB21							0.761
LB30						-0.431	0.804
LB28						0.424	0.755
LB29						0.399	0.793
LB12					0.339	-0.344	0.576

Note. 'Principal axis factoring' extraction method was used in combination with a 'promax' rotation

Table D.2. Factor Statistics Summary

Factor	SS Loadings	% of Variance	Cumulative %
1	2.77	9.22	9.22
2	2.52	8.39	17.61
3	2.18	7.28	24.89
4	2.07	6.91	31.80
5	1.46	4.85	36.65
6	1.16	3.86	40.51

Table D.3. Inter-Factor Correlations

	1	2	3	4	5	6
1	—	0.600	0.485	0.422	0.440	0.476
2		—	0.578	0.641	0.589	0.385
3			—	0.607	0.645	0.254
4				—	0.670	0.261
5					—	0.279
6						—

Table D.4. Model Fit Measures

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.0450	0.0360	0.0541	0.889	-1082	407	270	< .001

Appendix D Table D.5. Assumption Checks- Bartlett's Test of Sphericity

χ^2	df	p
2470	435	< .001

Appendix D Table D.6. KMO Measure of Sampling Adequacy

	MSA
Overall	0.859
ALB1	0.868
ALB2	0.861
ALB3	0.839
ALB4	0.905
ALB5	0.877
ALB6	0.864
ALB7	0.839
ALB8	0.848
ALB9	0.754
ALB10	0.815
ALB11	0.854
ALB12	0.823
ALB13	0.712
ALB14	0.868
ALB15	0.886
ALB16	0.887
ALB17	0.858
ALB18	0.901
ALB19	0.859
ALB20	0.894
ALB21	0.878
ALB22	0.941
ALB23	0.883
ALB24	0.685
ALB25	0.891
ALB26	0.847
ALB27	0.844
ALB28	0.792
ALB29	0.800

	MSA
ALB30	0.633

Appendix D Table D.6. *Eigenvalues*

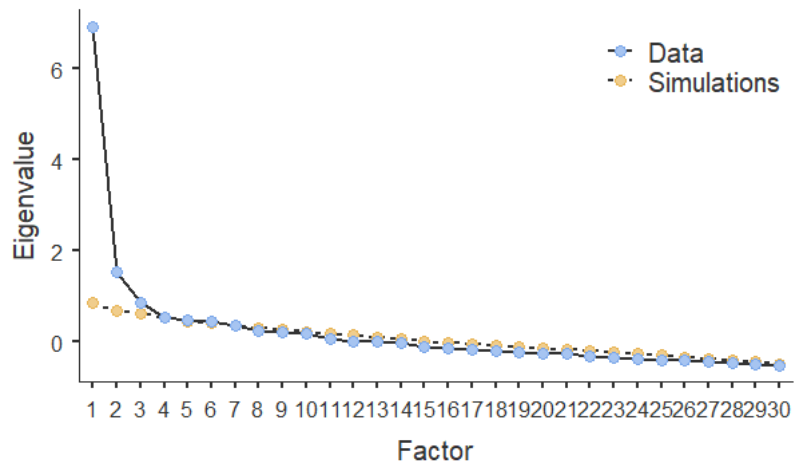
Factor	Eigenvalue
1	6.92935
2	1.55696
3	0.85350
4	0.54979
5	0.48275
6	0.45939
7	0.36284
8	0.23635
9	0.22619
10	0.18357
11	0.08562
12	0.02571
13	0.00997
14	-0.03000
15	-0.09084
16	-0.14618
17	-0.17309
18	-0.19039
19	-0.20810
20	-0.24795
21	-0.26174
22	-0.31892
23	-0.34285
24	-0.36912
25	-0.38500
26	-0.40868
27	-0.43240
28	-0.45057
29	-0.47431
30	-0.50222

Appendix D Table D.7. EFA Descriptives

	ALSA	ALSR	ALSE	ALB
N	248	245	241	249
Mean	5.40	5.20	72.1	4.95
Median	5.43	5.17	74.6	4.97
Mode	5.36	5.33	68.7 ^a	4.83
Standard deviation	0.547	0.693	15.4	0.599
Minimum	3.79	3.17	31.0	3.30
Maximum	6.86	7.00	100	6.63
Skewness	-0.332	-0.0990	-0.608	-0.156
Std. error skewness	0.155	0.156	0.157	0.154
Kurtosis	0.185	-0.127	-0.202	-0.171
Std. error kurtosis	0.308	0.310	0.312	0.307

^a More than one mode exists, only the first is reported

Appendix D Figure D.1. Scree Plot



Appendix E - CFA Results for ALB Scale | First Dataset (Model 1)

Appendix E Table E.1. Factor Loadings

Factor	Indicator	Estimate	SE	Z	p	Stand. Estimate
Factor 1	LB8	0.983	0.0672	14.62	< .001	0.719
	LB7	1.096	0.0580	18.90	< .001	0.859
	LB15	0.933	0.0519	17.97	< .001	0.830
	LB11	0.587	0.0491	11.96	< .001	0.620
	LB23	0.407	0.0473	8.60	< .001	0.469
	LB17	0.486	0.0528	9.20	< .001	0.497
	LB26	0.540	0.0615	8.78	< .001	0.477
Factor 2	LB2	0.681	0.0658	10.35	< .001	0.580
	LB3	0.818	0.0603	13.57	< .001	0.712
	LB4	0.769	0.0635	12.12	< .001	0.650
	LB5	0.962	0.0767	12.54	< .001	0.669
	LB6	0.890	0.0673	13.22	< .001	0.699
Factor 3	LB14	0.699	0.0584	11.97	< .001	0.804
	LB13	0.471	0.0570	8.27	< .001	0.505
	LB24	0.343	0.0713	4.81	< .001	0.296

Appendix E Table E.2. Factor Estimates/Covariances

		Estimate	SE	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a				
	Factor 2	0.602	0.0463	13.01	< .001	0.602
	Factor 3	0.691	0.0553	12.50	< .001	0.691
Factor 2	Factor 2	1.000 ^a				
	Factor 3	0.423	0.0651	6.50	< .001	0.423
Factor 3	Factor 3	1.000 ^a				

^a fixed parameter

Appendix E Table E.3. Test for Exact Fit

χ^2	df	p
275	87	<.001

Appendix E Table E.4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AIC	BIC
				Lower	Upper		
0.890	0.867	0.0569	0.0800	0.0694	0.0907	13965	14149

Appendix E Table E.5. CFA Dataset 1 Descriptives

	ALSA	ALSR	ALSE	ALB	NSA	SJS	MRM	MSE	TLB
N	347	344	333	337	349	338	344	338	337
Missing	3	6	17	13	1	12	6	12	13
Mean	5.56	5.37	72.8	5.02	5.14	4.79	5.11	66.6	5.83
Median	5.60	5.33	75.1	5.07	5.33	4.80	5.13	70.0	5.86
Mode	5.90	5.6	78.6	5.60	5.00	5.20	5.27 ^a	73.6 ^a	5.86
Sum	1928	1847	24256	1691	1795	1620	1758	22523	1965
Standard deviation	0.568	0.700	14.2	0.663	0.835	0.915	0.559	17.7	0.579
Minimum	3.90	3.17	25.0	3.00	1.33	1.60	3.60	18.1	3.86
Maximum	7.00	7.00	100	6.93	7.00	7.00	6.73	100	7.00
Skewness	-0.149	-0.164	-0.619	-0.195	-0.525	-0.304	-0.191	-0.517	-0.213
Std. error skewness	0.131	0.131	0.134	0.133	0.131	0.133	0.131	0.133	0.133
Kurtosis	-0.0938	-0.155	-0.0465	-0.128	1.00	0.323	0.0179	-0.285	-0.0587
Std. error kurtosis	0.261	0.262	0.266	0.265	0.260	0.265	0.262	0.265	0.265

Appendix F - CFA Results for ALB Scale | Dataset 1 (Model 2)

Appendix F Table F.1. Factor Loadings

Factor	Indicator	Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
Factor 1	LB8	0.998	0.0670	0.867	1.129	14.91	< .001	0.730
	LB7	1.117	0.0574	1.005	1.230	19.46	< .001	0.876
	LB15	0.946	0.0515	0.845	1.047	18.38	< .001	0.842
	LB11	0.571	0.0495	0.474	0.668	11.54	< .001	0.604
Factor 2	LB2	0.682	0.0658	0.553	0.811	10.36	< .001	0.581
	LB3	0.817	0.0604	0.699	0.936	13.53	< .001	0.711
	LB4	0.770	0.0635	0.645	0.894	12.12	< .001	0.650
	LB5	0.962	0.0767	0.812	1.113	12.54	< .001	0.670
	LB6	0.889	0.0673	0.757	1.021	13.21	< .001	0.698
Factor 3	LB14	0.693	0.0577	0.580	0.806	12.02	< .001	0.798
	LB13	0.474	0.0567	0.363	0.585	8.36	< .001	0.508
	LB24	0.346	0.0713	0.206	0.486	4.85	< .001	0.300
Factor 4	LB23	0.568	0.0499	0.470	0.666	11.39	< .001	0.654
	LB17	0.625	0.0566	0.514	0.736	11.05	< .001	0.639
	LB26	0.740	0.0651	0.612	0.867	11.36	< .001	0.654

Appendix F Table F.2. Factor Covariances

		95% Confidence Interval						
		Estimate	SE	Lower	Upper	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a						
	Factor 2	0.588	0.0474	0.495	0.681	12.42	< .001	0.588
	Factor 3	0.664	0.0555	0.555	0.773	11.97	< .001	0.664
	Factor 4	0.663	0.0486	0.568	0.759	13.66	< .001	0.663
Factor 2	Factor 2	1.000 ^a						
	Factor 3	0.424	0.0653	0.296	0.552	6.49	< .001	0.424
	Factor 4	0.457	0.0634	0.333	0.582	7.21	< .001	0.457
Factor 3	Factor 3	1.000 ^a						
	Factor 4	0.619	0.0679	0.485	0.752	9.11	< .001	0.619
Factor 4	Factor 4	1.000 ^a						

^a fixed parameter

Table F.3. Test for Exact Fit

χ^2	df	p
201	84	< .001

Appendix F Table F.4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AIC	BIC
				Lower	Upper		
0.931	0.914	0.0486	0.0644	0.0530	0.0758	13898	14093

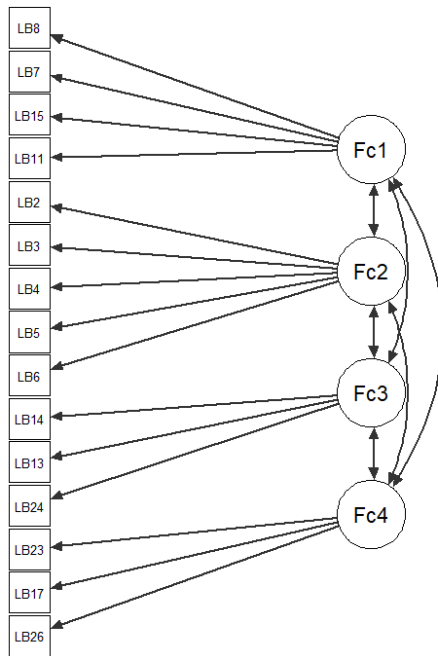
Table F.5. ALB Scale Reliability Statistics

Cronbach's α	
scale	0.858

Appendix G Table F.6. ALB Items' Reliability Statistics

	If item dropped
	Cronbach's α
LB2	0.852
LB3	0.848
LB4	0.849
LB5	0.853
LB6	0.847
LB11	0.845
LB8	0.843
LB13	0.858
LB14	0.849
LB15	0.842
LB7	0.840
LB17	0.853
LB23	0.852
LB24	0.864
LB26	0.850

Figure F.1. Path Diagram



Appendix G - CFA Results for ALB Scale | Dataset 2 (Model 1)

Appendix G Table G.1. Factor Loadings

Factor	Indicator	Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
Factor 1	LB8	0.984	0.0562	0.874	1.094	17.50	< .001	0.688
	LB7	0.910	0.0520	0.808	1.012	17.52	< .001	0.698
	LB15	0.828	0.0479	0.734	0.922	17.29	< .001	0.689
	LB11	0.738	0.0427	0.655	0.822	17.28	< .001	0.680
	LB26	0.669	0.0522	0.567	0.772	12.83	< .001	0.537
	LB23	0.505	0.0406	0.425	0.584	12.43	< .001	0.525
	LB17	0.501	0.0421	0.418	0.583	11.88	< .001	0.503
Factor 2	LB2	0.887	0.0552	0.779	0.995	16.07	< .001	0.648
	LB3	0.932	0.0526	0.829	1.035	17.74	< .001	0.698
	LB4	0.850	0.0509	0.750	0.949	16.70	< .001	0.664
	LB5	1.152	0.0569	1.040	1.263	20.23	< .001	0.769
	LB6	1.021	0.0555	0.912	1.129	18.38	< .001	0.716
Factor 3	LB13	0.408	0.0461	0.318	0.499	8.86	< .001	0.428
	LB14	0.700	0.0529	0.596	0.803	13.24	< .001	0.737
	LB24	0.327	0.0592	0.210	0.443	5.51	< .001	0.276

Appendix G Table G.2. Factor Covariances

		95% Confidence Interval						
		Estimate	SE	Lower	Upper	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a						
	Factor 2	0.640	0.0343	0.572	0.707	18.66	< .001	0.640
	Factor 3	0.725	0.0495	0.628	0.822	14.66	< .001	0.725
Factor 2	Factor 2	1.000 ^a						
	Factor 3	0.416	0.0553	0.308	0.524	7.53	< .001	0.416
Factor 3	Factor 3	1.000 ^a						

^a fixed parameter

Appendix G Table G.3. Test for Exact Fit

χ^2	df	p
435	87	< .001

Appendix G Table G.4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AIC	BIC
				Lower	Upper		
0.874	0.848	0.0552	0.0830	0.0753	0.0908	25562	25771

Appendix H - CFA Results for ALB Scale | Second Dataset (Model 2)

Appendix H Table H.1. Factor Loadings

Factor	Indicator	Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
Factor 1	LB8	1.022	0.0560	0.912	1.132	18.24	< .001	0.714
	LB7	1.011	0.0502	0.913	1.110	20.13	< .001	0.776
	LB15	0.899	0.0469	0.807	0.991	19.18	< .001	0.748
	LB11	0.697	0.0446	0.609	0.784	15.62	< .001	0.641
Factor 2	LB2	0.890	0.0552	0.781	0.998	16.12	< .001	0.650
	LB3	0.933	0.0526	0.830	1.036	17.74	< .001	0.699
	LB4	0.851	0.0509	0.751	0.951	16.72	< .001	0.665
	LB5	1.150	0.0570	1.038	1.261	20.15	< .001	0.768
	LB6	1.019	0.0556	0.910	1.128	18.33	< .001	0.715
Factor 3	LB13	0.401	0.0461	0.311	0.491	8.70	< .001	0.420
	LB14	0.684	0.0516	0.583	0.785	13.24	< .001	0.720
	LB24	0.358	0.0592	0.242	0.474	6.04	< .001	0.303
Factor 4	LB26	0.806	0.0539	0.700	0.911	14.96	< .001	0.647
	LB23	0.632	0.0415	0.551	0.714	15.22	< .001	0.658
	LB17	0.631	0.0432	0.546	0.715	14.60	< .001	0.633

Appendix H Table H.2. Factor Estimates/Variances

				95% Confidence Interval				
		Estimate	SE	Lower	Upper	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a						
	Factor 2	0.585	0.0382	0.510	0.660	15.33	< .001	0.585
	Factor 3	0.629	0.0509	0.529	0.729	12.37	< .001	0.629
	Factor 4	0.636	0.0412	0.555	0.717	15.42	< .001	0.636
Factor 2	Factor 2	1.000 ^a						
	Factor 3	0.425	0.0558	0.315	0.534	7.61	< .001	0.425
	Factor 4	0.523	0.0448	0.436	0.611	11.69	< .001	0.523
Factor 3	Factor 3	1.000 ^a						
	Factor 4	0.762	0.0567	0.651	0.873	13.45	< .001	0.762
Factor 4	Factor 4	1.000 ^a						

^a fixed parameter

Appendix H Table H3. Test for Exact Fit

χ^2	df	p
295	84	< .001

Appendix H Table H4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AIC	BIC
				Lower	Upper		
0.924	0.905	0.0481	0.0657	0.0577	0.0740	25428	25650

Appendix H Table H5. ALB Scale Reliability Statistics

Cronbach's α	
scale	0.860

Appendix H Table H6. ALB Items Reliability Statistics

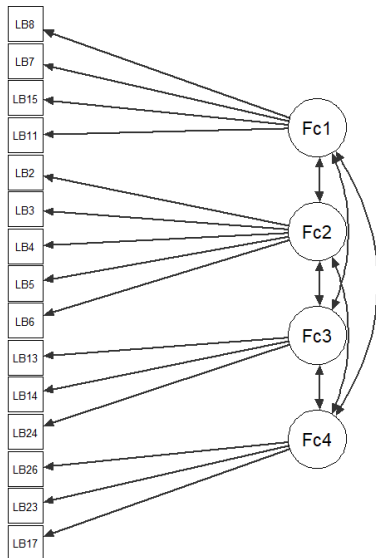
If item dropped	
Cronbach's α	
LB2	0.852
LB3	0.848
LB4	0.849
LB5	0.846
LB6	0.847
LB11	0.846
LB8	0.846
LB13	0.861
LB14	0.853
LB15	0.850
LB7	0.848
LB17	0.853
LB23	0.853
LB24	0.864
LB26	0.852

Appendix H Table H.7. CFA Dataset 2 Descriptives

	ALSA	ALSR	ALSE	ALB	MRM	MSE	SJS	LI	PCS	TLB
N	602	601	546	581	601	587	583	581	596	581
Mean	5.53	5.38	71.5	4.94	5.12	66.3	4.86	5.05	4.24	5.79
Median	5.60	5.50	74.0	5.00	5.13	70.6	5.00	5.25	4.25	5.86
Mode	5.70	5.33 ^a	77.1 ^a	4.93	5.20 ^a	75.7 ^a	5.20 ^a	6.00	4.00	6.00
Sum	3329	3236	39064	2867	3080	38910	2832	2934	2528	3364
Standard deviation	0.614	0.734	14.3	0.714	0.525	19.2	0.957	1.16	0.442	0.607
Minimum	3.20	3.17	29.6	3.00	3.47	5.14	2.20	1.00	3.00	4.00
Maximum	7.00	7.00	99.4	7.00	6.60	100	7.00	7.00	5.50	7.00
Skewness	-0.570	-0.364	-0.647	-0.270	-0.165	-0.774	-0.270	-0.747	0.337	-0.342
Kurtosis	0.602	-0.176	-0.0375	0.0416	-0.0653	0.179	-0.359	0.467	0.209	0.0673

^a More than one mode exists, only the first is reported

Appendix H Figure H1. Path Diagram



Appendix I - Convergent, Concurrent, Discriminant Validity

Appendix I Table I.1. Correlation Matrix in Dataset 1

		ALB	TLB	SJS
ALB	Pearson's r	—		
	df	—		
	p-value	—		
TLB	Pearson's r	0.587	—	
	df	333	—	
	p-value	< .001	—	
SJS	Pearson's r	0.110	0.313	—
	df	332	332	—
	p-value	0.044	< .001	—

Appendix I Table I.2. Correlation Matrix in Dataset 2

		ALB	TLB	LI	SJS
ALB	Pearson's r	—			
	df	—			
	p-value	—			
TLB	Pearson's r	0.632	—		
	df	576	—		
	p-value	< .001	—		
LI	Pearson's r	0.413	0.438	—	0.251
	df	575	575	—	577
	p-value	< .001	< .001	—	< .001
SJS	Pearson's r	0.196	0.340		—
	df	577	577		—
	p-value	< .001	< .001		—

Appendix J - Focus Group Session Design

Appendix J Table J.1. Focus Group 1 Session Design on Zoom

Sections	Time (Mins)	
Introduction of the project and process	10	Large Room Presentation
Group discussion on keeping the items	15	Small Group Discussion
Share out the group discussion	15 mins (3 minutes share-4 minutes feedback from other groups)	Large Room Presentation
Large group discussion on revising the items	45	Large Room Presentation
Closure	5	

Appendix J Table J.2. Focus Group 2 Session Design on Zoom

Section	Time (Mins)	
Introduction & Instructions	10	Large Room Presentation
Label all scales in groups	30	Small Group Discussion
Label Presentations and arguments	15	Large Room Presentation
Reflections, questions, comments on Group's Label presentations	10	Large Room Discussion
Finalizing the labels	25	Large Room Discussion
Closure	5	

Appendix K - EFA Results for ALSE Scale

Appendix K Table K.1. Factor Loadings

	Factor		Uniqueness
	1	2	
SE3	0.882		0.314
SE1	0.812		0.354
SE2	0.738		0.439
SE10	0.563		0.454
SE6	0.483	0.443	0.321
SE5		0.909	0.252
SE9		0.800	0.378
SE7	0.323	0.543	0.396
SE8	0.439	0.475	0.338
SE4	0.317	0.396	0.596

Note. 'Principal axis factoring' extraction method was used in combination with a 'oblimin' rotation

Appendix K Table K.2. Factor Statistics Summary

Factor	SS Loadings	% of Variance	Cumulative %
1	3.34	33.4	33.4
2	2.82	28.2	61.6

Appendix K Table K.3. Inter-Factor Correlations

	1	2
1	—	0.584
2		—

Appendix K Table K.4. Model Fit Measures

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.107	0.0849	0.131	0.912	-45.7	96.3	26	< .001

Table K.5. Bartlett's Test of Sphericity

χ^2	df	p
1436	45	< .001

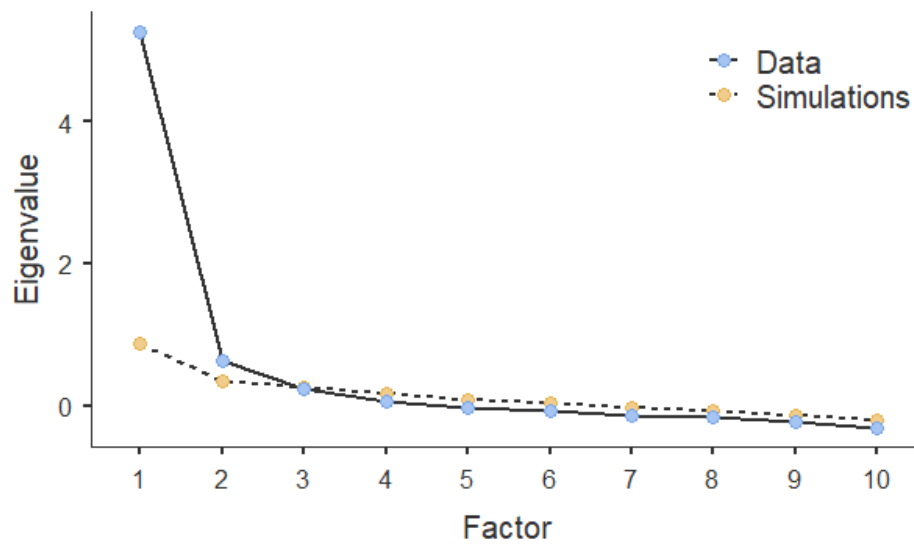
Appendix K Table K.6. KMO Measure of Sampling Adequacy

	MSA
Overall	0.902
ALSE1	0.919
ALSE2	0.900
ALSE3	0.897
ALSE4	0.913
ALSE5	0.845
ALSE6	0.919
ALSE7	0.942
ALSE8	0.918
ALSE9	0.858
ALSE10	0.905

Appendix K Table K.5. Initial Eigenvalues

Factor	Eigenvalue
1	5.2625
2	0.6311
3	0.2263
4	0.0611
5	-0.0241
6	-0.0629
7	-0.1358
8	-0.1679
9	-0.2222
10	-0.3057

Appendix K Figure K.1. Scree Plot



Appendix L - CFA Results for ALSE Scale| First Dataset

Appendix L Table L.1. Factor Loadings

Factor	Indicator	Estimate	SE	Z	p	Stand. Estimate
Factor 1	ALSE1	13.8	0.814	17.0	< .001	0.828
	ALSE2	14.0	0.959	14.6	< .001	0.736
	ALSE3	14.9	0.907	16.4	< .001	0.808
Factor 2	ALSE5	15.9	0.893	17.8	< .001	0.852
	ALSE9	17.1	0.986	17.4	< .001	0.835
	ALSE7	12.9	1.046	12.3	< .001	0.648

Appendix L Table L.2. Factor Estimates/Covariances

		Estimate	SE	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a				
	Factor 2	0.660	0.0424	15.5	< .001	0.660
Factor 2	Factor 2	1.000 ^a				

^a fixed parameter

Appendix L Table L.3. Test for Exact Fit

χ^2	df	p
21.7	8	0.005

Appendix L Table L.4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AIC	BIC
				Lower	Upper		
0.985	0.971	0.0340	0.0718	0.0363	0.109	16531	16603

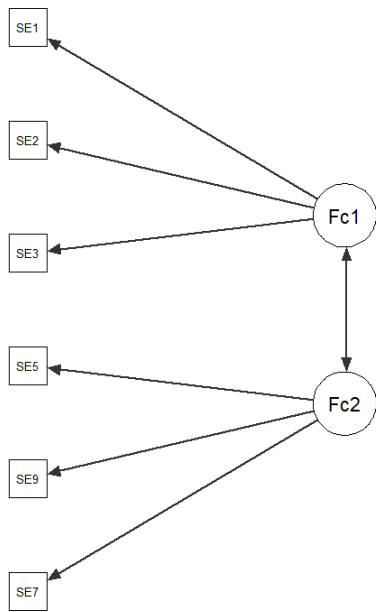
Appendix L Table L.5. Scale Reliability Statistics

Cronbach's α	
scale	0.854

Table L.6. Item Reliability Statistics

	If item dropped
	Cronbach's α
ALSE3	0.824
ALSE5	0.824
ALSE7	0.838
ALSE9	0.829
ALSE1	0.825
ALSE2	0.839

Appendix L Figure L.1. Path Diagram



Appendix M - CFA Results for ALSE Scale | Second Dataset

Appendix M Table M.1. Factor Loadings

Factor	Indicator	Estimate	SE	Z	p	Stand. Estimate
Factor 1	SE1	12.7	0.666	19.0	< .001	0.748
	SE2	14.7	0.732	20.2	< .001	0.779
	SE3	15.9	0.711	22.3	< .001	0.846
Factor 2	SE5	16.1	0.707	22.7	< .001	0.841
	SE9	18.3	0.767	23.9	< .001	0.873
	SE7	14.7	0.852	17.3	< .001	0.689

Appendix M Table M.2. Factor Estimates/Covariances

		Estimate	SE	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a				
	Factor 2	0.635	0.0339	18.7	< .001	0.635
Factor 2	Factor 2	1.000 ^a				

^a fixed parameter

Appendix M Table M.3. Test for Exact Fit

χ^2	df	p
30.3	8	< .001

Appendix M Table M.4. Fit Measures Comparisons

Models (Dataset 2)	CFI	TLI	SRMR	RMSEA	RMSEA 90% CI	
					Lower	Upper
Final Model (without ALSE 10 & 4)	0.986	0.973	0.0307	0.0712	0.0454	0.0991
EFA Model with ASLE 10	0.971	0.954	0.0445	0.0857	0.0657	0.107
EFA Model with ASLE 4	0.947	0.914	0.0411	0.116	0.0963	0.136
EFA Model with ASLE 10 & 4	0.934	0.903	0.0525	0.115	0.0989	0.132

Appendix M Table M.5. Scale Reliability Statistics

Cronbach's α	
scale	0.875

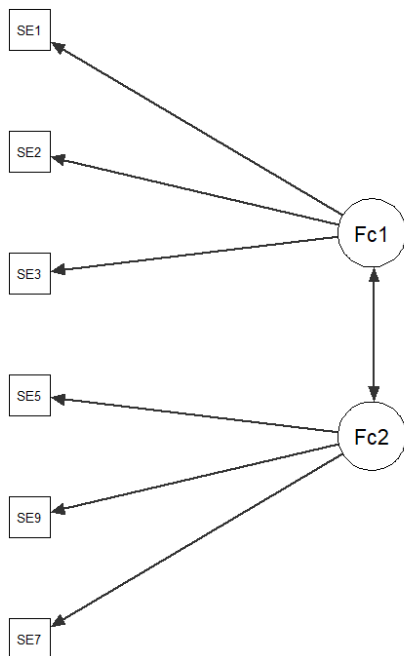
Appendix M Table M.6. Item Reliability Statistics

Item-rest correlation	
SE1	0.641
SE2	0.616
SE3	0.693
SE5	0.687
SE9	0.678
SE7	0.627
SE10	0.657

Appendix M Table M.7. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI	
				Lower	Upper
0.986	0.973	0.0307	0.0712	0.0454	0.0991

Appendix M Figure M.1. Path Diagram



Appendix N - Correlation Matrix (Convergent-Concurrent-Discriminant Validity)

Appendix N Table N.1. Correlation Matrix (Dataset 1)

		ALSA	SA-Sister	ALSE	MSE	ALB	TLB	MRM	SJS
ALSA	Pearson's r	—							
	df	—							
	p-value	—							
SA-Sister	Pearson's r	0.552	—						
	df	345	—						
	p-value	< .001	—						
ALSE	Pearson's r	0.506	0.494	—					
	df	326	327	—					
	p-value	< .001	< .001	—					
MSE	Pearson's r	0.456	0.469	0.796	—				
	df	333	334	327	—				
	p-value	< .001	< .001	< .001	—				
ALB	Pearson's r	0.434	0.292	0.470	0.471	—			
	df	332	333	325	332	—			
	p-value	< .001	< .001	< .001	< .001	—			
TLB	Pearson's r	0.555	0.360	0.505	0.491	0.587	—		
	df	332	333	325	332	333	—		
	p-value	< .001	< .001	< .001	< .001	< .001	—		
MRM	Pearson's r	0.274	0.109	0.268	0.199	0.344	0.287	—	
	df	339	340	328	335	334	334	—	
	p-value	< .001	0.044	< .001	< .001	< .001	< .001	—	
SJS	Pearson's r	0.293	0.323	0.251	0.220	0.109	0.320	0.030	—
	df	332	333	325	332	331	331	334	—
	p-value	< .001	< .001	< .001	< .001	0.048	< .001	0.586	—

Appendix N Table N.2. Correlation Matrix (Dataset 2)

		ALSR	MRM	ALSE	MSE	ALB	TLB	SJS	PC
ALSA	Pearson's r	0.407	0.232	0.559		0.537	0.547	0.223	0.101
	df	591	590	540		572	572	573	586
	p-value	< .001	< .001	< .001		< .001	< .001	< .001	0.014
	df	-2	-2	-2		-2	-2	-2	-2
ALSR	Pearson's r	—							
	df	—							
	p-value	—							
MRM	Pearson's r	0.344	—						
	df	597	—						
	p-value	< .001	—						
ALSE	Pearson's r	0.377	0.178	—					
	df	544	543	—					
	p-value	< .001	< .001	—					
MSE	Pearson's r	0.343	0.245	0.777	—	0.483	0.526	0.185	0.024
	df	581	581	542	—	573	572	574	577
	p-value	< .001	< .001	< .001	—	< .001	< .001	< .001	0.558
ALB	Pearson's r	0.532	0.290	0.476		—			
	df	578	578	538		—			
	p-value	< .001	< .001	< .001		—			
TLB	Pearson's r	0.413	0.304	0.495		0.607	—		
	df	579	578	537		576	—		
	p-value	< .001	< .001	< .001		< .001	—		
SJS	Pearson's r	0.233	0.076	0.186		0.192	0.340	—	
	df	580	580	537		577	577	—	
	p-value	< .001	0.068	< .001		< .001	< .001	—	
PC	Pearson's r	0.078	0.081	-0.020		0.058	0.086	-0.040	—
	df	592	592	540		575	575	576	—
	p-value	0.058	0.049	0.636		0.161	0.039	0.333	—

Appendix O - EFA Results for ALSA Scale

Appendix O Table O.1. Factor Loadings

	Factor			Uniqueness
	1	2	3	
SA11	0.718			0.472
SA1	0.610			0.530
SA10	0.445			0.719
SA2	0.415			0.720
SA12				0.922
SA13				0.952
SA14				0.967
SA8		0.702		0.503
SA6		0.589		0.547
SA7		0.541		0.619
SA9		0.480		0.629
SA4				0.905
SA3			0.679	0.489
SA5			0.399	0.693

Note. 'Maximum likelihood' extraction method was used in combination with a 'oblimin' rotation

Appendix O Table O.2. Factor Statistics Summary

Factor	SS Loadings	% of Variance	Cumulative %
1	1.64	11.71	11.7
2	1.61	11.53	23.2
3	1.08	7.70	31.0

Appendix O Table O.3. Inter-Factor Correlations

	1	2	3
1	—	0.337	0.205
2		—	0.338
3			—

Appendix O Table O.4. Model Fit Measures

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.0427	0.0185	0.0631	0.918	-211	75.6	52	0.018

Appendix O Table O.5. Bartlett's Test of Sphericity

χ^2	df	p
598	91	<.001

Appendix O Table O.6. KMO Measure of Sampling Adequacy

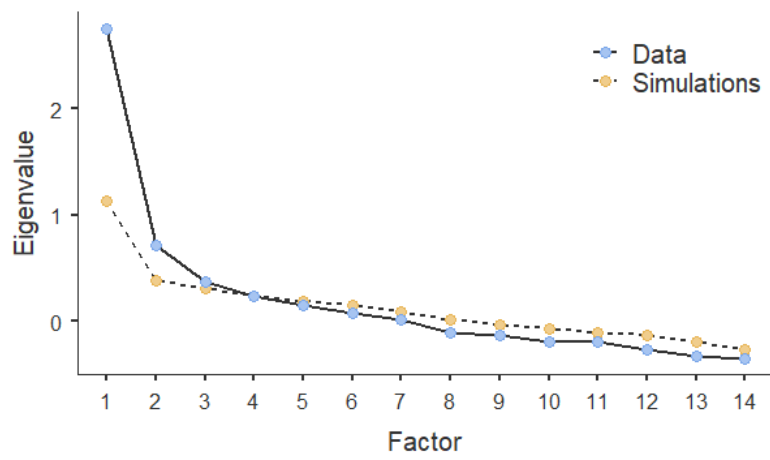
	MSA
Overall	0.775
SA1	0.766
SA2	0.809
SA3	0.753
SA4	0.687
SA5	0.785
SA6	0.763
SA7	0.818
SA8	0.810
SA9	0.786
SA10	0.799
SA11	0.737
SA12	0.800
SA13	0.676

MSA	
SA14	0.703

Appendix O Table O.7. Initial Eigenvalues

Factor	Eigenvalue
1	2.7557
2	0.7132
3	0.3711
4	0.2374
5	0.1498
6	0.0761
7	0.0196
8	-0.1090
9	-0.1305
10	-0.1897
11	-0.1987
12	-0.2641
13	-0.3267
14	-0.3487

Figure O.1. Scree Plot



Appendix P - CFA Results for ALSA Scale | First Dataset

Appendix P Table P.1. Factor Loadings

Factor	Indicator	Estimate	SE	Z	p	Stand. Estimate
Factor 1	SA1	0.696	0.0678	10.27	< .001	0.656
	SA2	0.428	0.0602	7.11	< .001	0.459
	SA10	0.545	0.0768	7.10	< .001	0.460
	SA11	0.518	0.0626	8.27	< .001	0.519
Factor 2	SA6	0.602	0.0519	11.60	< .001	0.636
	SA7	0.597	0.0457	13.09	< .001	0.705
	SA8	0.541	0.0423	12.81	< .001	0.693
	SA9	0.585	0.0581	10.07	< .001	0.565
Factor 3	SA3	0.697	0.0933	7.47	< .001	0.619
	SA5	0.508	0.0814	6.24	< .001	0.434

Appendix P Table P.2. Factor Estimates/Covariances

		Estimate	SE	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a				
	Factor 2	0.622	0.0644	9.66	< .001	0.622
	Factor 3	0.651	0.0943	6.90	< .001	0.651
Factor 2	Factor 2	1.000 ^a				
	Factor 3	0.641	0.0895	7.16	< .001	0.641
Factor 3	Factor 3	1.000 ^a				

^a fixed parameter

Appendix P Table P.3. Test for Exact Fit

χ^2	df	p
85.8	32	< .001

Appendix P Table P.4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI	
				Lower	Upper
0.915	0.880	0.0519	0.0696	0.0520	0.0876

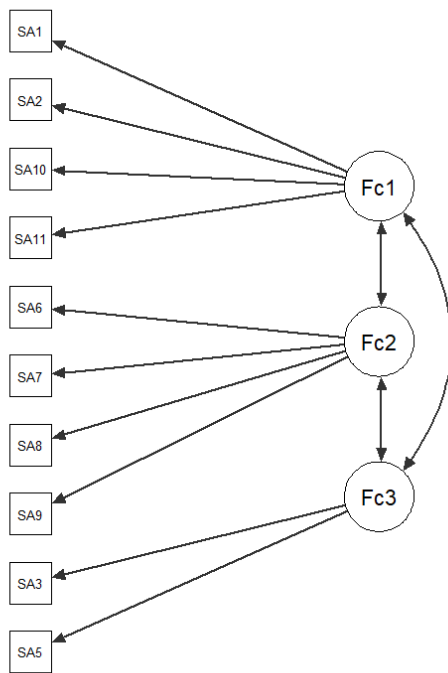
Appendix P Table P.5. Scale Reliability Statistics

Cronbach's α	
scale	0.755

Appendix P Table P.6. Item Reliability Statistics

	If item dropped
	Cronbach's α
SA1	0.728
SA2	0.740
SA3	0.735
SA5	0.756
SA6	0.726
SA7	0.721
SA8	0.723
SA9	0.731
SA10	0.743
SA11	0.743

Figure P.1. Path Diagram



Appendix Q - CFA Results for ALSA Scale | Second Dataset

Appendix Q Table Q.1. Factor Loadings

Factor	Indicator	Estimate	SE	Z	p	Stand. Estimate
Factor 1	SA1	0.667	0.0577	11.56	< .001	0.562
	SA11	0.442	0.0471	9.39	< .001	0.456
	SA10	0.566	0.0577	9.81	< .001	0.483
	SA2	0.663	0.0575	11.53	< .001	0.567
Factor 2	SA8	0.465	0.0363	12.80	< .001	0.571
	SA7	0.634	0.0430	14.74	< .001	0.654
	SA6	0.564	0.0442	12.77	< .001	0.569
	SA9	0.641	0.0519	12.35	< .001	0.560
Factor 3	SA3	0.945	0.0650	14.54	< .001	0.775
	SA5	0.790	0.0631	12.51	< .001	0.613

Appendix Q Table Q.2. Factor Covariances

		Estimate	SE	Z	p	Stand. Estimate
Factor 1	Factor 1	1.000 ^a				
	Factor 2	0.611	0.0538	11.34	< .001	0.611
	Factor 3	0.551	0.0551	9.99	< .001	0.551
Factor 2	Factor 2	1.000 ^a				
	Factor 3	0.585	0.0513	11.41	< .001	0.585
Factor 3	Factor 3	1.000 ^a				

^a fixed parameter

Appendix Q Table Q.3. Test for Exact Fit

χ^2	df	p
94.2	32	< .001

Appendix Q Table Q.4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI	
				Lower	Upper
0.937	0.912	0.0406	0.0568	0.0438	0.0703

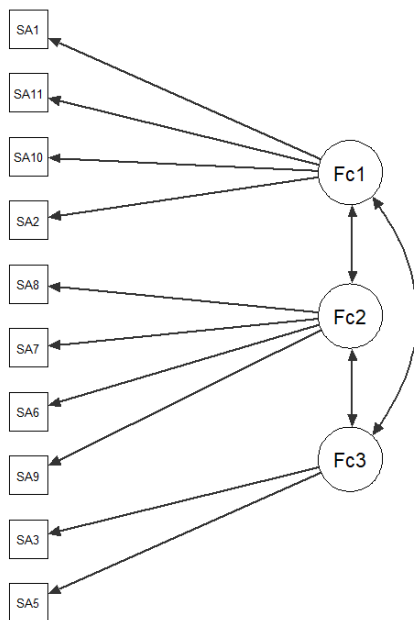
Appendix Q Table Q.4. Scale Reliability Statistics

Cronbach's α	
scale	0.752

Appendix Q Table Q.5. Scale Reliability Statistics

	If item dropped
	Cronbach's α
SA1	0.737
SA2	0.731
SA5	0.737
SA6	0.732
SA7	0.726
SA8	0.736
SA9	0.723
SA10	0.737
SA11	0.742
SA3	0.720

Appendix Q Figure Q.1. Path Diagram



Appendix R - EFA Results for ALSR Scale

Appendix R Table R.1. Factor Loadings

	Factor		Uniqueness
	1	2	
SR1	0.820		0.391
SR2	0.614		0.654
SR3	0.420		0.837
SR6	0.291	0.233	0.809
SR7 - Reverse Score	0.215	0.212	0.873
SR4 - Reverse Score	-0.212	0.705	0.574
SR5 - Reverse Score		0.424	0.823

Note. 'Maximum likelihood' extraction method was used in combination with a 'promax' rotation

Appendix R Table R.2. Factor Statistics Summary

Factor	SS Loadings	% of Variance	Cumulative %
1	1.320	18.9	18.9
2	0.718	10.3	29.1

Appendix R Table R.3. Inter-Factor Correlations

	1	2
1	—	0.387
2		—

Appendix R Table R.4. Model Fit Measures

RMSEA	RMSEA 90% CI		TLI	BIC	Model Test		
	Lower	Upper			χ^2	df	p
0.0630	0.00914	0.109	0.862	-28.2	15.8	8	0.045

Appendix R Table R.5. Bartlett's Test of Sphericity

χ^2	df	p
171	21	<.001

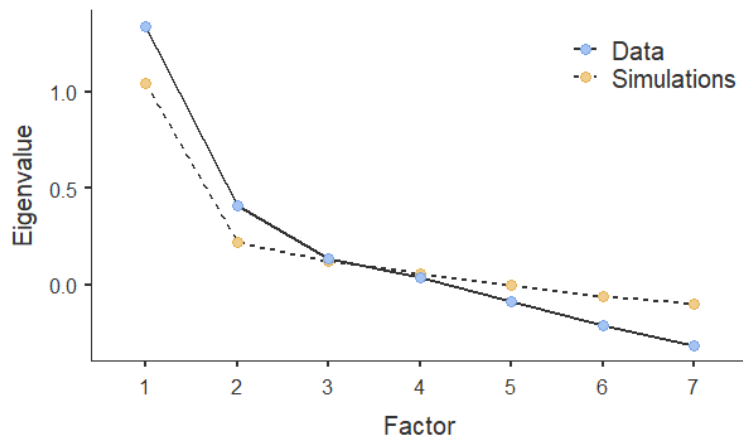
Appendix R Table R.6. KMO Measure of Sampling Adequacy

	MSA
Overall	0.636
SR1	0.622
SR2	0.657
SR3	0.704
SR4 - Reverse Score	0.518
SR7 - Reverse Score	0.685
SR6	0.650
SR5 - Reverse Score	0.583

Appendix R Table R.7. Initial Eigenvalues

Factor	Eigenvalue
1	1.3384
2	0.4109
3	0.1375
4	0.0401
5	-0.0822
6	-0.2062
7	-0.3102

Table R.7. Scree Plot



Appendix S - CFA Results for ALSR Scale

Appendix S Table S.1. Factor Loadings

Factor	Indicator	Estimate	SE	95% Confidence Interval		Z	p	Stand. Estimate
				Lower	Upper			
Factor 1	ASR1	0.896	0.0640	0.771	1.021	14.01	< .001	0.703
	ALSR2	0.973	0.0768	0.822	1.123	12.67	< .001	0.613
	ALSR3	0.502	0.0522	0.399	0.604	9.61	< .001	0.461
	ALSR7®	0.592	0.0657	0.464	0.721	9.01	< .001	0.442
Factor 2	ALSR4®	0.718	0.1927	0.341	1.096	3.73	< .001	0.509
	ALSR5®	0.635	0.1671	0.308	0.963	3.80	< .001	0.653

Appendix S Table S.2. Factor Covariances

		Estimate	SE	Z	p
Factor 1	Factor 1	1.000 ^a			
	Factor 2	0.198	0.0711	2.78	0.005
Factor 2	Factor 2	1.000 ^a			

^a fixed parameter

Appendix S Table S.3. Test for Exact Fit

χ^2	df	p
31.0	8	< .001

Appendix S Table S.4. Fit Measures

CFI	TLI	SRMR	RMSEA	RMSEA 90% CI		AIC	BIC
				Lower	Upper		
0.940	0.887	0.0416	0.0692	0.0446	0.0958	11571	11655

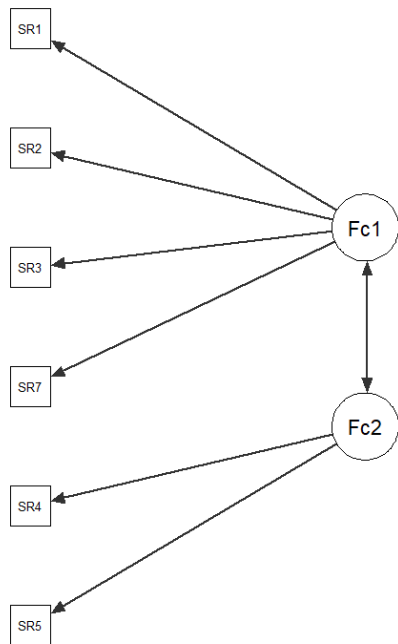
Appendix S Table S.5. Scale Reliability Statistics

Cronbach's α	
scale	0.574

Appendix S Table S.6. Item Reliability Statistics

	If item dropped
	Cronbach's α
SR1	0.475
SR2	0.513
SR3	0.530
SR4_Reverse_Score	0.592
SR5_Reverse	0.561
SR7_Reverse	0.488

Appendix S Table S.5. Path Diagram



Appendix T - Generalizability Test

Appendix T Table T.1. Model Fit Comparison Across Different Race, Gender, Sector, and Time Groups

Group	Scale	Number of participants (N)	Chi-Square			Model Fit						Reliability	Model Modification
			X ²	df	p	CFI	TLI	SRMR	RMSEA	AIC	BIC	Cronbach's α	
URM	ALB	132	136	84	<0.001	0.917	0.896	0.069	0.0682	5658	5805	0.848	
	ALSE	129	16.8	8	<0.05	0.974	0.952	0.0403	0.0926	6362	6416	0.835	
	ALSA	132	54.2	32	<0.01	0.910	0.874	0.0613	0.0724	3459	3554	0.744	
	ALSR												
Non-URM	ALB	592	305	84	<0.001	0.915	0.894	0.0524	0.0677	25391	25614	0.843	
	ALSE	585	35.3	8	<0.001	0.984	0.970	0.0328	0.0764	29111	29194	0.866	
	ALSA	613	61	19	<0.001	0.942	0.914	0.0399	0.06	13296	13406	0.739	SA5 with or without works
	ALSR	602	8.23	2	0.016	0.978	0.933	0.0214	0.719	7639	7692	0.624	Factor 2 removal improves model fit significantly.
Business	ALB	83	98.2	71	<0.05	0.936	0.918	0.07	0.068	3222	3338	0.85	ALB24 removed due to near zero loading
	ALSE	85	37.4	13	<0.001	0.905	0.846	0.0486	0.138	4234	4280	0.787	
	ALSA		94.6	32	<0.001	0.934	0.907	0.0414	0.0565	17026	17172		Factor structure significantly changed: Factor 1: SA7,8,11 /Factor2: SA1,2,3,5/ Factor3: SA 9, 10
	ALSR	88	11.1	8	<0.198	0.935	0.878	0.0679	0.0661	1701	1748	0.595	With SR7 in the second factor has a better model fit
Nonprofit	ALB	188	155	84	<0.001	0.910	0.888	0.0619	0.0670	8118	8283	0.847	
	ALSE	189	16.2	8	<0.05	0.986	0.973	0.0405	0.0737	9533	9595	0.862	
	ALSA	188	47.1	32	<0.05	0.947	0.926	0.0549	0.0496	5318	5426	0.689	
	ALSR	191	23.6	8	<0.01	0.835	0.690	0.101	0.068	3658	3720	0.5	With SR7 in the second factor has a better model fit
Government	ALB	140	125	84	<0.01	0.921	0.901	0.0655	0.0587	6088	6238	0.817	
	ALSE	131	15.1	8	<0.05	0.966	0.945	0.0589	0.085	7643	7706	0.831	Include SE10 improves power
	ALSA	140	46.3	32	<0.05	0.891	0.847	0.0676	0.0565	3966	4063	0.618	Removing SA3 and SA10 improves the model
	ALSR	4.67	8	0.792	0.792	1	1.09	0.0361	0	2727	2783	0.552	With SR7 in the second factor has a better model fit
Education/Ma	ALB	191	152	84	<0.001	0.930	0.913	0.0667	0.0656	8221	8387	0.902	
	ALSE	185	25.7	13	<0.05	0.984	0.974	0.0315	0.0725	10618	10688	0.909	Include SE10 improves the model significantly

Women	ALSA	191	67.8	32	<0.001	0.898	0.857	0.0610	0.0765	5323	5431	0.758	Adding SE10 weakens the model significantly
	ALSR												
	ALB	566	323	84	<0.001	0.910	0.888	0.0557	0.0708	24474	24695	0.848	
	ALSE	558	37.7	8	<0.001	0.981	0.965	0.0324	0.0816	27905	27987	0.863	
	ALSA	581	107	32	<0.001	0.924	0.894	0.0458	0.0637	15944	16088	0.747	
	ALSR	572	6.97	2	<0.05	0.980	0.940	0.021	0.0658	7316	7368	0.614	
Men	ALB	181	156	84	<0.001	0.919	0.899	0.056	0.0688	7592	7755	0.869	SE10 improves the model fit and α significantly!
	ALSE	173	29.5	13	<0.01	0.967	0.947	0.0518	0.0857	9815	9884	0.850	
	ALSA	186	46.3	32	<0.05	0.935	0.909	0.0521	0.049	5219	5326	0.706	
	ALSR	185	17.8	8	<0.05	0.931	0.871	0.0518	0.0812	3582	3643	0.608	
Millennials	ALB	326	205	84	<0.001	0.917	0.896	0.0572	0.0666	13955	14148	0.836	
	ALSE	315	24.4	8	<0.01	0.977	0.957	0.0368	0.0808	15503	15575	0.833	
	ALSA	337	62.1	32	<0.01	0.942	0.919	0.0446	0.0529	9216	9342	0.731	
	ALSR	332	24.9	8	<0.01	0.922	0.854	0.0407	0.0796	6281	6353	0.566	
Gen X	ALB	238	165	84	<0.001	0.933	0.916	0.0495	0.638	10118	10295	0.877	Removing SA2 improves the model
	ALSE	229	24	8	<0.01	0.978	0.96	0.0504	0.0937	11250	11315	0.873	
	ALSA	242	45.1	24	<0.01	0.937	0.906	0.0462	0.0603	5808	5913	0.706	
	ALSR	Error											
Gen Z	ALB	94	119	71	<0.001	0.902	0.874	0.0728	0.0837	4208	4331	0.848	ALB removal improves model. It is negatively correlated with the scale.
	ALSE	93	7.27	8	<0.001	0.929	0.909	0.0575	0.104	7898	7974	0.907	One factor model works
	ALSA	96	52.5	32	<0.147	0.948	0.932	0.0702	0.514	2684	27.63	0.815	Factor structure significantly changed: (Factor 1: SA6,7,8,10,11 Factor2: SA1,2,3,5,9)
	ALSR												Model did not converge
	ALB	182	199	84	<0.001	0.899	0.874	0.0681	0.0868	6228	6991	0.886	
T3 (30 days after)	ALSE	187	18.2	8	<0.05	0.986	0.973	0.0452	0.0826	8642	8704	0.893	SE10 should be removed!
	ALSA	206	73.3	32	<0.001	0.904	0.865	0.0630	0.0791	4592	4701	0.780	
	ALSR												
	ALSE	224											
	ALSA	239	99.4	32	<0.001	0.862	0.806	0.0625	0.0939	6790	5905	0.762	
	ALSR	232											

Appendix U - Regression for Predictive Validity

Appendix U.1- Predictive Validity of ALB (T1) on LI (T3)

Appendix U Table U.1.1. Linear Regression Model Fit Measures

Model	R	R ²	Adjusted R ²	AIC	BIC	RMSE	Overall Model Test			
							F	df1	df2	p
1	0.463	0.214	0.209	1022	1188	13.4	39.3	1	144	< .001

Appendix U Table U.1.2. Model Coefficients - T3-LI – Box-Cox

Predictor	Estimate	SE	t	p
Intercept	-16.6	8.38	-1.98	0.05
ALB	10.5	1.68	6.27	< .001

Appendix U Table U.1.3. Durbin–Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
-0.0707	2.12	0.490

Appendix U Table U.1.4. Collinearity Statistics

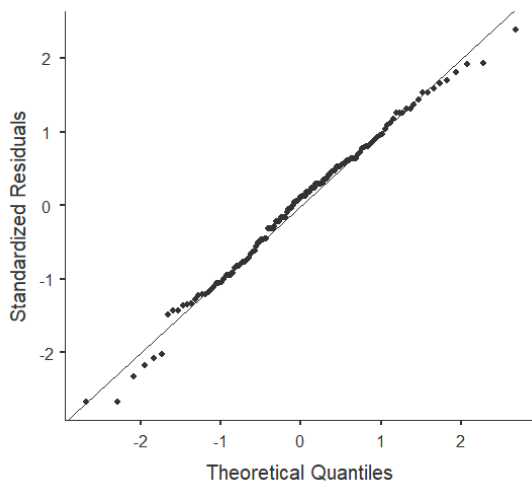
	VIF	Tolerance
ALB	1.00	1.00

Appendix U Table U.1.5. Normality Test (Shapiro-Wilk)

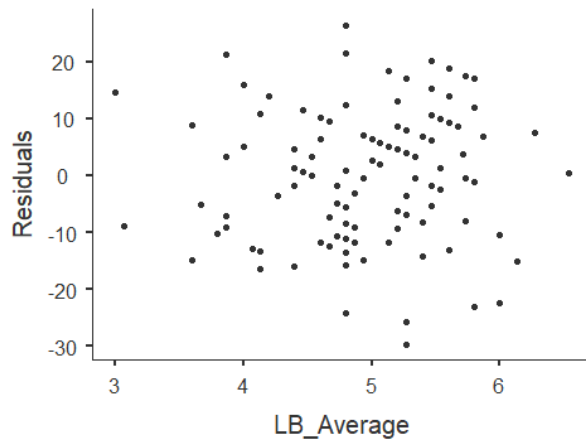
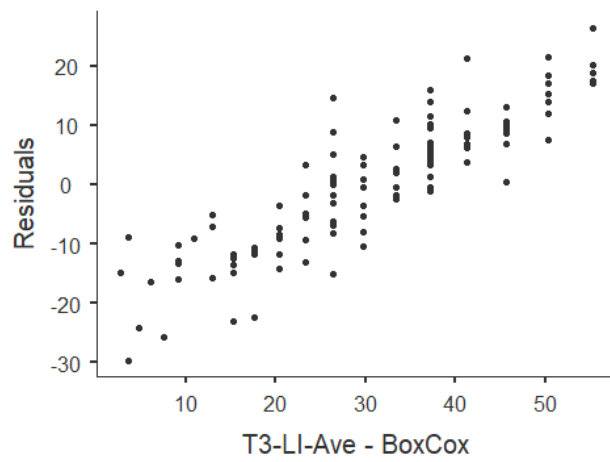
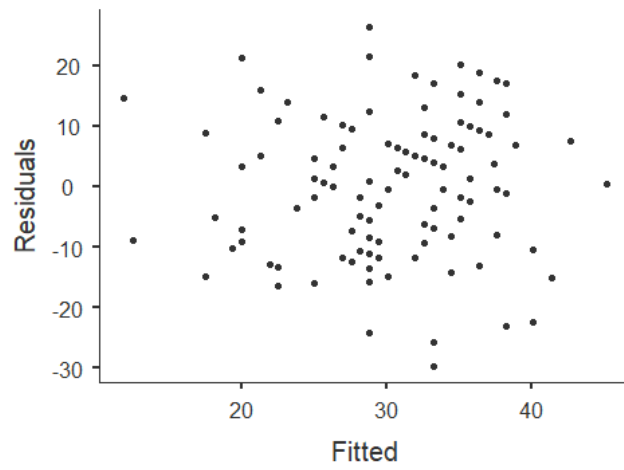
Statistic	p
0.990	0.400

Appendix U Table U.1.6. Predictive Validity Descriptives for ALB

	LI-T3	ALB-T1	LI-T3 - BoxCox
N	146	146	146
Mean	5.38	4.95	35.4
Median	5.75	5.03	39.2
Mode	6.00	5.27	43.9
Sum	786	723	5173
Standard deviation	1.04	0.670	15.2
Minimum	2.25	3.00	2.88
Maximum	7.00	6.53	66.4
Skewness	-0.985	-0.561	-0.158
Std. error skewness	0.201	0.201	0.201
Kurtosis	0.719	0.341	-0.428
Std. error kurtosis	0.399	0.399	0.399

Appendix U Figure U.1.1. Q-Q Plot

Appendix U Figure U.1.1. Residuals Plots



Appendix U.2- Predictive Validity of ALSE (T1) on ALB (T3)

Table U.2.1. Linear Regression Model Fit Measures

Model	R	R ²	Adjusted R ²	AIC	BIC	RMSE	Overall Model Test			
							F	df1	df2	p
1	0.288	0.0831	0.0769	281	290	0.604	13.4	1	148	< .001

Appendix U Table U.2.2. Model Coefficients - T3-ALB

Predictor	Estimate	SE	t	p
Intercept	4.4321	0.24862	17.83	< .001
ALSE	0.0121	0.00330	3.66	< .001

Appendix U Table U.2.3. Durbin–Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
0.103	1.76	0.156

Appendix U Table U.2.4. Collinearity Statistics

	VIF	Tolerance
ALSE	1.00	1.00

Appendix U Table U.2.4. Normality Test (Shapiro-Wilk)

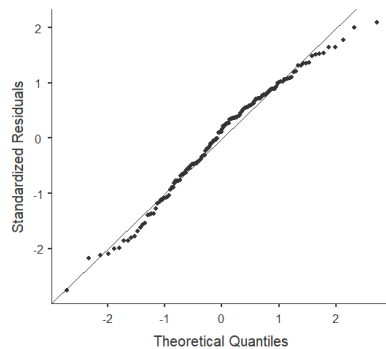
Statistic	p
0.982	0.050

Table U.2.6. Predictive Validity Test Descriptives for ALSE, ALSA, & ALSR

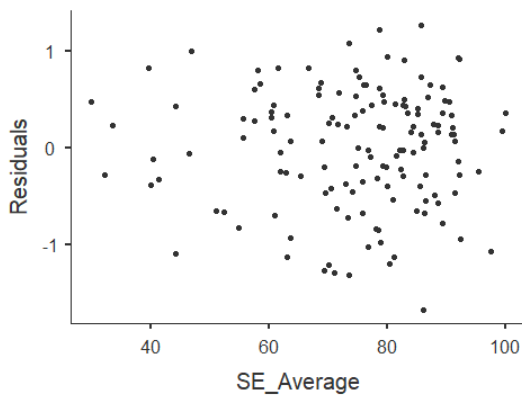
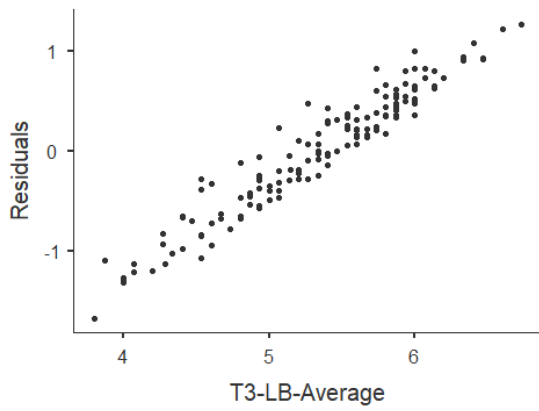
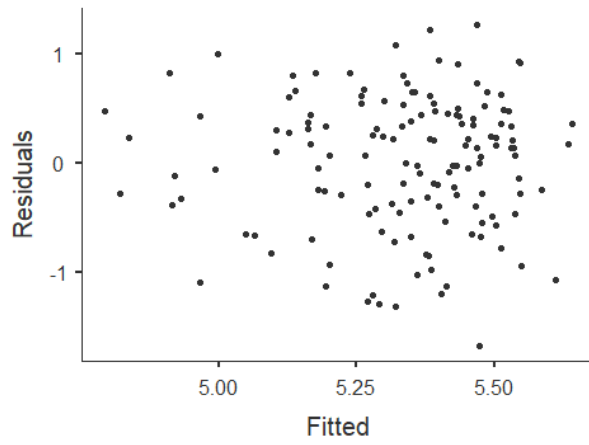
	TLB-T3	ALSE-T1	ALB-T3	ALSE- BoxCox	ALSA-T1	ALSR-T1
N	150	150	150	150	150	150
Missing	0	0	0	0	0	0
Mean	5.95	73.8	5.32	19135	5.50	5.38
Median	6.00	76.7	5.40	19599	5.60	5.33
Mode	6.00	78.6 ^a	5.87 ^a	20817 ^a	5.90	5.83
Sum	893	11069	799	2.87e+6	826	807
Standard deviation	0.635	15.1	0.633	8189	0.689	0.734
Minimum	4.00	30.0	3.80	1898	3.00	3.33
Maximum	7.00	100	6.73	37925	7.00	6.83
Skewness	-0.634	-0.872	-0.341	-0.181	-0.861	-0.450
Std. error skewness	0.198	0.198	0.198	0.198	0.198	0.198
Kurtosis	0.349	0.289	-0.512	-0.615	1.53	0.0433
Std. error kurtosis	0.394	0.394	0.394	0.394	0.394	0.394

^a More than one mode exists, only the first is reported

Appendix U Figure U.2.1. Q-Q Plot



Appendix U Figure U.2.2. Q-Q Plot Residuals Plots



Appendix U.3- Predictive Validity of ALSA (T1) on ALB (T3)

Appendix U Table U.3.1. Linear Regression Model Fit Measures

Model	R	R ²	Adjusted R ²	AIC	BIC	RMSE	Overall Model Test			
							F	df1	df2	p
1	0.287	0.0822	0.0760	281	290	0.605	13.3	1	148	<.001

Appendix U Table U.3.2. Model Coefficients - T3-ALB

Predictor	Estimate	SE	t	p
Intercept	3.872	0.4018	9.64	<.001
ALSA	0.264	0.0724	3.64	<.001

Appendix U Table U.3.3. Durbin–Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
0.0562	1.86	0.396

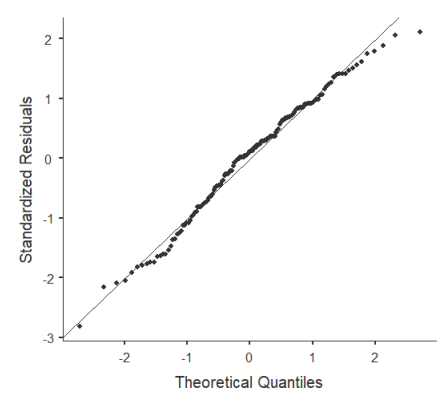
Appendix U Table U.3.4. Collinearity Statistics

	VIF	Tolerance
ALSA	1.00	1.00

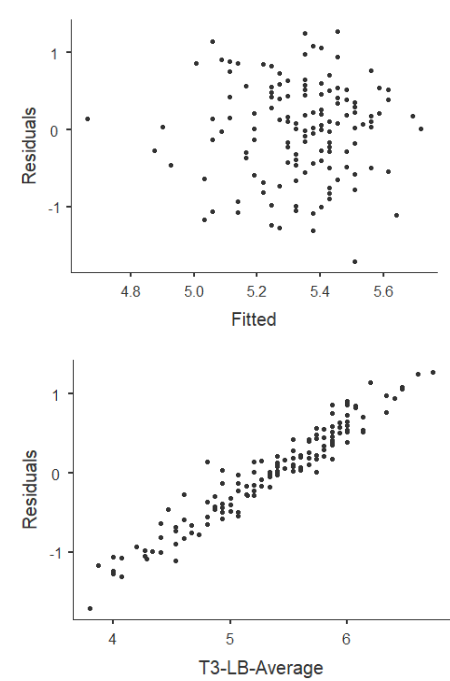
Appendix U Table U.3.5. Normality Test (Shapiro-Wilk)

Statistic	p
0.986	0.125

Appendix U Figure U.3.1. Q-Q Plot



Appendix U Figure U.3.1. Residuals Plots



Appendix U.4- Predictive Validity of ALSR (T1) on ALB (T3)

Appendix U Table U.4.1. Linear Regression Model Fit Measures

Model	R	R ²	Adjusted R ²	AIC	BIC	RMSE	Overall Model Test			
							F	df1	df2	p
1	0.317	0.101	0.0946	278	287	0.599	16.6	1	148	< .001

Appendix U Table U.4.2. Model Coefficients - T3-ALB-Average

Predictor	Estimate	SE	t	p
Intercept	3.851	0.3652	10.54	< .001
ALSR	0.274	0.0673	4.07	< .001

Appendix U Table U.4.3. Durbin–Watson Test for Autocorrelation

Autocorrelation	DW Statistic	p
0.0564	1.87	0.440

Appendix U Table U.4.4. Collinearity Statistics

	VIF	Tolerance
ALSR	1.00	1.00

Appendix U Table U.4.5. Normality Test (Shapiro-Wilk)

Statistic	p
0.984	0.086

Appendix U Figure U.4.1. Q-Q Plot

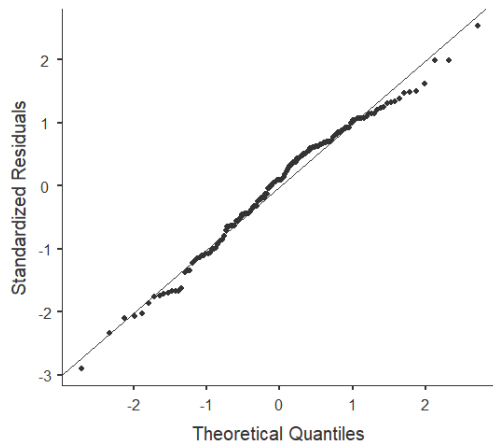
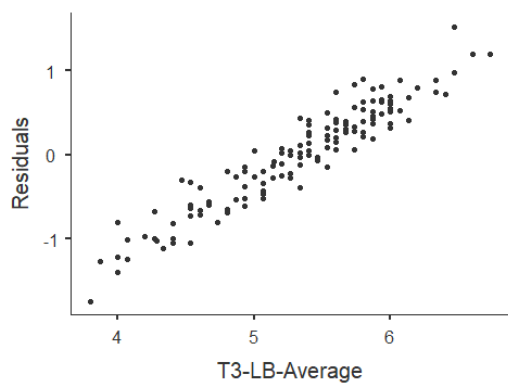
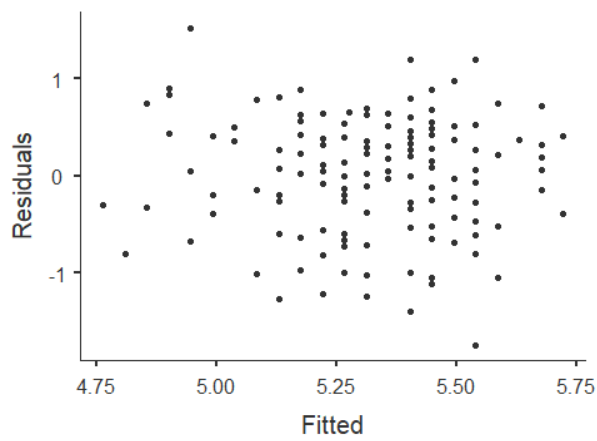


Figure U.4.2. Residuals Plots



Appendix V - LMM Models

Appendix V.1- Null Model 1 LMM Results

Appendix V Table V.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 +(1 ID)
AIC	686.337
BIC	698.205
LogLikel.	-340.169
R-squared Marginal	0.000
R-squared Conditional	0.589
Converged	yes
Optimizer	bobyqa

Appendix V Table V.2. Fixed Effects Parameter Estimates

Names	Estimate	SE	95% Confidence Interval		df	t	p
			Lower	Upper			
(Intercept)	5.22	0.0466	5.12	5.31	157	112	< .001

Appendix V Table V.3. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.515	0.266	0.589
Residual		0.430	0.185	

Note. Number of Obs: 386, groups: ID 160

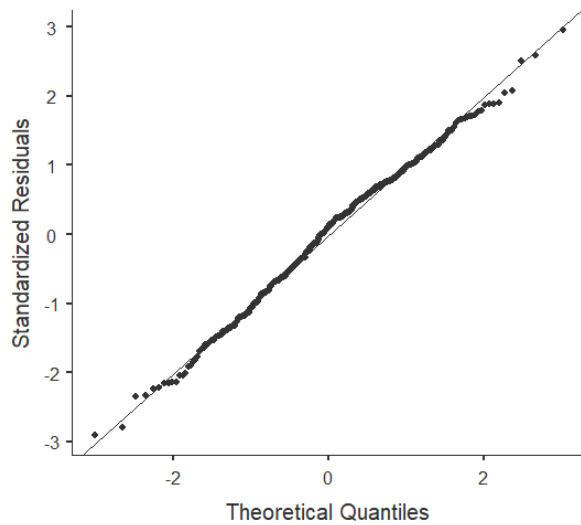
Appendix V Table V.4. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	2.00	789	105	1.00	< .001

Appendix V Table V.5. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0556	0.184

Appendix V Figure V.1. Q-Q Plot



Appendix V.2- Model 10 LMM Results

Appendix V Table V.2.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 + MRM + ALSR + ALSA + ALSE + MRM: ALSE+(1 + MRM:ALSE ID)
AIC	524.987
BIC	564.468
LogLikel.	-252.494
R-squared Marginal	0.355
R-squared Conditional	0.736
Converged	yes
Optimizer	bobyqa

Appendix V Table V.2.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
MRM	9.34	1	279.8	0.002
ALSR	28.42	1	358.9	< .001
ALSA	31.80	1	326.0	< .001
ALSE	13.33	1	279.7	< .001
MRM * ALSE	6.83	1	34.3	0.013

Note. Satterthwaite method for degrees of freedom

Table V.2.3. Fixed Effects Parameter Estimates

Names	Estimate	SE	95% Confidence Interval		df	t	p
			Lower	Upper			
(Intercept)	5.21251	0.03852	5.13701	5.2880	147.3	135.31	< .001
MRM	0.18625	0.06095	0.06679	0.3057	279.8	3.06	0.002
ALSR	0.22003	0.04128	0.13913	0.3009	358.9	5.33	< .001
ALSA	0.29887	0.05300	0.19500	0.4028	326.0	5.64	< .001
ALSE	0.00838	0.00230	0.00388	0.0129	279.7	3.65	< .001
MRM * ALSE	0.00956	0.00366	0.00239	0.0167	34.3	2.61	0.013

Appendix V Table V.2.4. Random Components

Random Components				
Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.41906	0.176	0.597
	MRM * ALSE	0.00997	9.95e-5	
Residual		0.34419	0.118	

Note. Number of Obs: 383, groups: ID 160

Appendix V Table V.2.5. Random Parameters Correlations

Groups	Param.1	Param.2	Corr.
ID	(Intercept)	MRM*ALSE	0.795

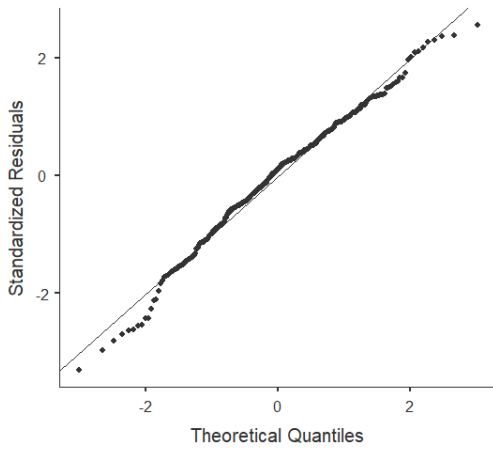
Appendix V Table V.2.6. Random Effect LRT

Test	N. par	AI C	LRT	df	p
MRM*ALSE in (1 + MRM*ALSE ID)	8	525	4.34	2.00	0.114

Appendix V Table V.2.7. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0531	0.230

Appendix V Figure V.2.1. Q-Q Plot



Appendix V Table V.2.8. Descriptives

	Time	ID	ALSA	ALSR	MRM	ALSE	ALB
N	1	160	159	159	159	158	157
	3	160	159	155	154	152	146
	4	84	84	84	84	84	83
Missing	1	0	1	1	1	2	3
	3	0	1	5	6	8	14
	4	0	0	0	0	0	1
Mean	1	83.3	5.57	5.41	5.17	73.7	5.05
	3	83.3	5.93	5.59	5.28	79.0	5.34
	4	76.8	5.94	5.60	5.24	79.4	5.37
Median	1	83.5	5.60	5.50	5.20	77.1	5.08
	3	83.5	6.00	5.60	5.27	81.7	5.42
	4	75.5	6.00	5.60	5.20	82.1	5.42
Mode	1	1.00 ^a	5.90	5.83	5.20	76.0 ^a	5.25
	3	151	6.00	5.60	5.27	76.1 ^a	5.50 ^a
	4	2.00 ^a	6.00	5.20 ^a	5.07	77.1 ^a	5.25
Sum	1	13323	886	860	822	11650	793
	3	13324	943	867	812	12003	780
	4	6447	499	470	440	6671	445
Standard deviation	1	47.1	0.576	0.741	0.471	15.4	0.647
	3	47.1	0.478	0.694	0.515	14.1	0.688
	4	39.0	0.467	0.737	0.485	15.3	0.610
Minimum	1	1	4.00	3.33	3.80	22.7	3.33
	3	1	4.60	3.40	3.93	29.3	3.17
	4	2	4.30	3.60	4.40	33.9	3.75
Maximum	1	166	7.00	7.00	6.60	100	6.92
	3	166	7.00	7.00	6.53	100	6.75
	4	143	6.80	7.00	6.87	100	6.67
Skewness	1	-0.00613	-0.194	-0.428	0.0338	-1.09	-0.186
	3	-0.00571	-0.298	-0.443	0.123	-1.11	-0.547
	4	-0.0203	-0.666	-0.562	0.673	-1.00	-0.304
Kurtosis	1	-1.17	-0.187	0.00354	-0.153	1.05	0.0990
	3	-1.17	0.0545	0.0788	0.00340	1.27	0.131
	4	-1.15	1.37	0.188	0.949	0.657	0.107

^a More than one mode exists, only the first is reported

Appendix W - YLE Impact on Competencies

Appendix W.1- Impact on ALB

Appendix W Table W.1.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 + Time+(1 ID)
AIC	646.3993
BIC	666.1785
LogLikel.	-318.1996
R-squared Marginal	0.0450
R-squared Conditional	0.6587
Converged	yes
Optimizer	bobyqa

Appendix W Table W.1.2. Fixed Effect Omnibus Tests

	F	Num df	Den df	p
Time	24.1	2	237	< .001

Note. Satterthwaite method for degrees of freedom

Appendix W Table W.1.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.245	0.0470	5.153	5.3376	165	111.66	< .001
Time1	linear	0.216	0.0401	0.137	0.2942	242	5.37	< .001
Time2	quadratic	-0.104	0.0356	-0.174	-0.0343	235	-2.92	0.004

Appendix W Table W.1.4. Post Hoc Comparisons - Time

Comparison							
Time	Time	Difference	SE	t	df	p _{bonferroni}	
1	- 3	-0.2798	0.0459	-6.092	234	< .001	
1	- 4	-0.3048	0.0570	-5.344	246	< .001	
3	- 4	-0.0250	0.0579	-0.432	246	1.000	

Appendix W Table W.1.5. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.211	0.0497	5.113	5.3081	175	104.94	< .001
Time1	linear	0.242	0.0412	0.161	0.3228	245	5.88	< .001
Time2	quadratic	-0.101	0.0356	-0.171	-0.0313	238	-2.84	0.005

Appendix W Table W.1.6. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.526	0.276	0.643
Residual		0.392	0.154	

Note. Number of Obs: 386 , groups: ID 160

Appendix W Table W.1.7. Random Effect LRT

Test	N. par	AI C	LRT	df	p
(1 ID)	4.00	774	129	1.00	< .001

Appendix W Table W.1.8. Post Hoc Comparisons – Time

Comparison			Difference	SE	t	df	p _{bonferroni}
Time		Time					
1	-	3	-0.2798	0.0459	-6.092	234	< .001
1	-	4	-0.3048	0.0570	-5.344	246	< .001
3	-	4	-0.0250	0.0579	-0.432	246	1.000

Appendix W Table W.1.9. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0401	0.565

Appendix W.2- Impact on ALSA

Appendix W Table W.2.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALSA ~ 1 + Time+(1 ID)
AIC	576.383
BIC	596.608
LogLikel.	-283.191
R-squared Marginal	0.104
R-squared Conditional	0.594
Converged	yes
Optimizer	bobyqa

Appendix W Table W.2.2. Fixed Effect Omnibus Tests

	F	Num df	Den df	p
Time	54.1	2	281	< .001

Note. Satterthwaite method for degrees of freedom

Appendix W Table W.2.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.802	0.0384	5.727	5.878	141	151.28	< .001
Time1	linear	0.253	0.0309	0.192	0.313	281	8.19	< .001
Time2	quadratic	-0.199	0.0310	-0.259	-0.138	281	-6.41	< .001

Appendix W Table W.2.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.403	0.162	0.547
Residual		0.367	0.134	

Note. Number of Obs: 422, groups: ID 141

Appendix W Table W.2.4. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	4.00	693	119	1.00	< .001

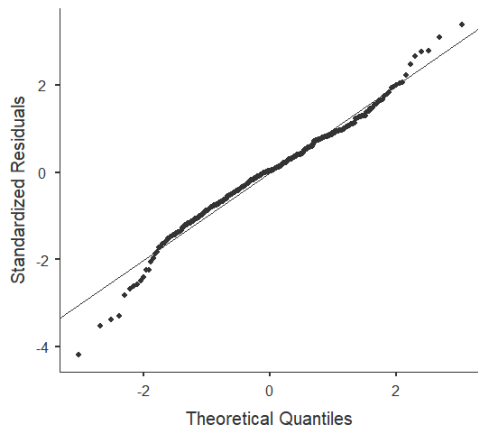
Appendix W Table W.2.4. Post Hoc Comparisons - Time

Comparison							
Time	Time	Difference	SE	t	df	pbonferroni	
1	- 2	-0.4220	0.0439	-9.61	283	< .001	
1	- 3	-0.3574	0.0438	-8.16	283	< .001	
2	- 3	0.0645	0.0439	1.47	283	0.429	

Appendix W Table W.2.5. Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0449	0.362

Appendix W Figure W.2.1. Q-Q Plot



Appendix W.3- Impact on MRM

Appendix W Table W.3.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	MRM ~ 1 + Time+(1 ID)
AIC	438.42340
BIC	458.63660
LogLikel.	-214.21170
R-squared Marginal	0.00892
R-squared Conditional	0.70080
Converged	yes
Optimizer	bobyqa

Appendix W Table W.3.2. Fixed Effect Omnibus Tests

	F	Num df	Den df	p
Time	6.25	2	280	0.002

Note. Satterthwaite method for degrees of freedom

Appendix W Table W.3.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.2100	0.0390	5.1335	5.2865	141	133.521	< .001
Time1	2 - 1	0.1174	0.0341	0.0506	0.1842	281	3.445	< .001
Time2	3 - (1, 2)	0.0226	0.0294	-0.0350	0.0803	280	0.769	0.443

Appendix W Table W.3.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.433	0.1875	0.698
Residual		0.285	0.0811	

Note. Number of Obs: 421, groups: ID 141

Appendix W Table W.3.5. Random Effect LRT

Test	N. par	AI C	LRT	df	p
(1 ID)	4.00	650	214	1.00	< .001

Appendix W Table W.3.6. Post Hoc Comparisons - Time

Comparison							
Time		Time	Difference	SE	t	df	p _{bonferroni}
1	-	2	-0.1174	0.0342	-3.43	282	0.002
1	-	3	-0.0813	0.0340	-2.39	282	0.053
2	-	3	0.0361	0.0342	1.06	282	0.877

Appendix W Table W.3.7. Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0516	0.213

Appendix W.4- Impact on ALSR

Appendix W Table W.4.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALSR ~ 1 + Time+(1 ID)
AIC	812.0364
BIC	832.2496
LogLikel.	-401.0182
R-squared Marginal	0.0284
R-squared Conditional	0.5954
Converged	yes
Optimizer	bobyqa

Appendix W Table W.4.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
Time	14.7	2	280	< .001

Note. Satterthwaite method for degrees of freedom

Appendix W Table W.4.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.5498	0.0529	5.4462	5.653	141	104.995	< .001
Time1	2 - 1	0.3044	0.0570	0.1927	0.416	280	5.341	< .001
Time2	3 - (1, 2)	0.0439	0.0492	-0.0525	0.140	280	0.892	0.373

Appendix W Table W.4.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.564	0.318	0.584
Residual		0.476	0.227	

Note. Number of Obs: 421 , groups: ID 141

Appendix W Table W.4.5. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	4.00	947	137	1.00	< .001

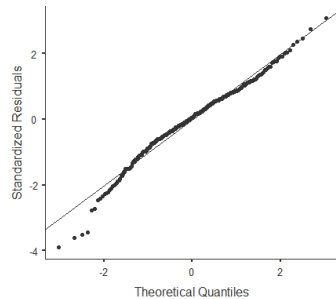
Appendix W Table W.4.6. Post Hoc Comparisons - Time

Comparison							
Time	Time	Difference	SE	t	df	pbonferroni	
1	- 2	-0.304	0.0572	-5.32	283	< .001	
1	- 3	-0.196	0.0569	-3.45	282	0.002	
2	- 3	0.108	0.0572	1.89	283	0.178	

Appendix W Table W.4.7. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0568	0.132

Appendix W Figure W.4.1. Q-Q Plot



Appendix W.5- Impact on ALSE

Appendix W Table W.5.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALSE ~ 1 + Time+(1 ID)
AIC	3259.3227
BIC	3279.4035
LogLikel.	-1624.6614
R-squared Marginal	0.0158
R-squared Conditional	0.6817
Converged	yes
Optimizer	bobyqa

Appendix W Table W.5.1. Fixed Effect Omnibus Tests

	F	Num df	Den df	p
Time	10.1	2	269	< .001

Note. Satterthwaite method for degrees of freedom

Appendix W Table W.5.2. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	76.28	1.206	73.913	78.64	136	63.24	< .001
Time1	2 - 1	4.28	1.108	2.109	6.45	269	3.86	< .001
Time2	3 - (1, 2)	2.18	0.958	0.304	4.06	269	2.28	0.024

Appendix W Table W.5.3. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	13.19	173.9	0.677
Residual		9.12	83.1	

Note. Number of Obs: 410 , groups: ID 139

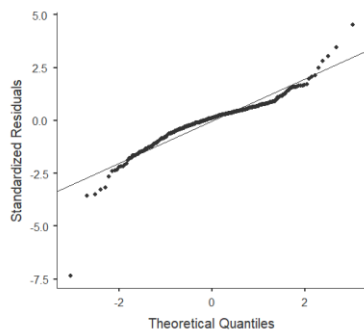
Appendix W Table W.5.4. Post Hoc Comparisons - Time

Comparison							
Time	Time	Difference	SE	t	df	pbonferroni	
1	- 2	-4.2802	1.11	-3.8497	274	< .001	
1	- 3	-4.3209	1.11	-3.8965	274	< .001	
2	- 3	-0.0407	1.11	-0.0366	274	1.000	

Appendix W Table W.5.5. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0985	< .001
Shapiro-Wilk	0.9013	< .001

Appendix W Figure W.5.1. Q-Q Plot



Appendix W.6. Descriptives

Appendix W Table W.6.1. YLE Evaluation Descriptive for ALSE, ALSR, & ALSA

	Time	ALSR	ALSA	MRM	ALSE	ALSE_SQRT
N	1	141	141	141	137	137
	2	139	140	139	136	136
	3	141	141	141	137	137
Missing	1	0	0	0	4	4
	2	2	1	2	5	5
	3	0	0	0	4	4
Mean	1	5.38	5.54	5.14	74.0	8.55
	2	5.68	5.96	5.26	77.7	8.77
	3	5.58	5.90	5.23	78.0	8.74
Median	1	5.40	5.60	5.20	77.1	8.78
	2	5.80	6.00	5.27	80.9	9.00
	3	5.60	5.90	5.20	81.4	9.02
Minimum	1	3.00	3.90	3.73	22.7	4.77
	2	3.40	4.50	3.53	25.1	5.01
	3	3.40	4.00	3.60	0.00	0.00
Maximum	1	7.00	7.00	6.13	100	10.0
	2	7.00	7.00	6.47	99.4	9.97
	3	7.00	7.00	6.53	100	10.0
Skewness	1	-0.491	-0.237	-0.375	-1.16	-1.54
	2	-0.593	-0.271	-0.233	-1.22	-1.65
	3	-0.441	-0.742	-0.206	-1.66	-3.44
Kurtosis	1	0.155	-0.0581	0.132	1.25	2.57
	2	0.315	-0.170	0.226	1.58	3.36
	3	-0.0367	1.42	0.524	3.68	19.0

Appendix X - LFC Impact on Competencies

Appendix X.1- LFC Impact on ALB

Appendix X Table X.1.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 + Time+(1 ID)
AIC	279.8718
BIC	292.6436
LogLikel.	-135.9359
R-squared Marginal	0.0763
R-squared Conditional	0.6115
Converged	yes
Optimizer	bobyqa

Appendix X Table X.1.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
Time	35.0	1	90.6	< .001

Note. Satterthwaite method for degrees of freedom

Appendix X Table X.1.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.313	0.0535	5.208	5.417	90.1	99.34	< .001
Time1	4 - 1	0.327	0.0553	0.219	0.436	90.6	5.92	< .001

Appendix X Table X.1.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.434	0.189	0.579
Residual		0.370	0.137	

Note. Number of Obs: 180, groups: ID 90

Appendix X Table X.1.5. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	3.00	315	36.9	1.00	< .001

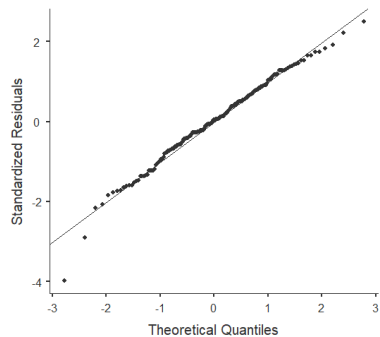
Appendix X Table X.1.6. Post Hoc Comparisons - Time

Comparison						
Time	Time	Difference	SE	t	df	p _{bonferroni}
1	- 4	-0.327	0.0556	-5.88	91.5	< .001

Appendix X Table X.1.7. Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0453	0.853

Appendix X Figure X.1.1. Q-Q Plot



Appendix X.2- LFC Impact on ALSE

Appendix X Table X.2.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALSE ~ 1 + Time+(1 ID)
AIC	1373.3291
BIC	1386.1009
LogLikel.	-682.6646
R-squared Marginal	0.0368
R-squared Conditional	0.7557
Converged	yes
Optimizer	bobyqa

Appendix X Table X.2.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
Time	26.8	1	90.4	< .001

Note. Satterthwaite method for degrees of freedom

Appendix X Table X.2.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	75.41	1.298	72.87	77.95	90.1	58.12	< .001
Time1	4 - 1	5.13	0.992	3.19	7.08	90.4	5.18	< .001

Appendix X Table X.2.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	11.38	129.4	0.746
Residual		6.63	44.0	

Note. Number of Obs: 180 , groups: ID 90

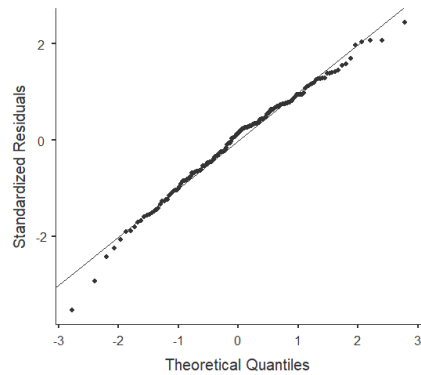
Appendix X Table X.2.5. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	3.00	1445	73.5	1.00	< .001

Appendix X Table X.2.6. Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0681	0.374

Appendix X Figure X.2.1. Q-Q Plot



Appendix X.3- LFC Impact on ALSA

Appendix X Table X.3.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALSA ~ 1 + Time+(1 ID)
AIC	245.368
BIC	258.140
LogLikel.	-118.684
R-squared Marginal	0.107
R-squared Conditional	0.586
Converged	yes
Optimizer	bobyqa

Appendix X Table X.3.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
Time	46.1	1	90.4	< .001

Note. Satterthwaite method for degrees of freedom

Appendix X Table X.3.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.821	0.0471	5.729	5.913	89.9	123.62	< .001
Time1	4 - 1	0.352	0.0518	0.250	0.453	90.4	6.79	< .001

Appendix X Table X.3.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.373	0.139	0.536
Residual		0.347	0.120	

Note. Number of Obs: 180, groups: ID 90

Appendix X Table X.3.5. Random Effect LRT

Test	N. par	AI C	LRT	df	p
(1 ID)	3.00	274	30.3	1.00	< .001

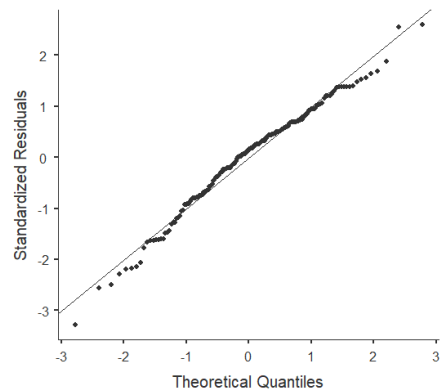
Appendix X Table X.3.6. Post Hoc Comparisons - Time

Comparison		Difference	SE	t	df	p _{bonferroni}
Time	Time					
1	- 4	-0.352	0.0521	-6.75	91.5	< .001

Appendix X Table X.3.7. Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0720	0.309

Appendix X Figure X.3.1. Q-Q Plot



Appendix X.4- Impact on MRM

Appendix X Table X.4.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	<code>MRM ~ 1 + Time+(1 ID)</code>
AIC	221.23950
BIC	234.01130
LogLikel.	-106.61974
R-squared Marginal	0.00205
R-squared Conditional	0.78049
Converged	yes
Optimizer	bobyqa

Appendix X Table X.4.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
Time	1.66	1	90.3	0.200

Note. Satterthwaite method for degrees of freedom

Appendix X Table X.4.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.2324	0.0551	5.1245	5.340	90.1	95.04	< .001
Time1	4 - 1	0.0501	0.0388	-0.0260	0.126	90.3	1.29	0.200

Appendix X Table X.4.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.489	0.2389	0.780
Residual		0.260	0.0674	

Note. Number of Obs: 180 , groups: ID 90

Appendix X Table X.4.5. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	3.00	304	84.5	1.00	< .001

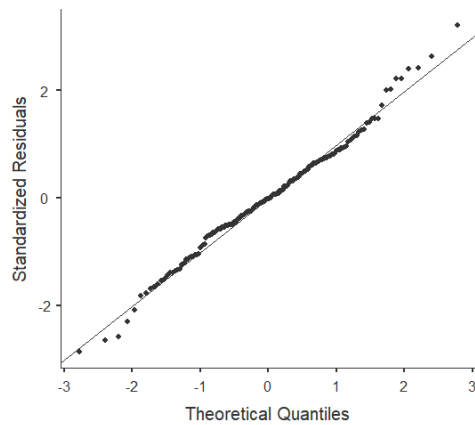
Appendix X Table X.4.6. Post Hoc Comparisons - Time

Comparison			Difference	SE	t	df	p _{bonferroni}
Time	Time						
1	-	4	-0.0501	0.0390	-1.28	91.3	0.203

Appendix X Table X.4.7. Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0569	0.604

Appendix X Figure X.4.1. Q-Q Plot



Appendix X.5- Impact on ALSR

Appendix X Table X.5.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALSR ~ 1 + Time+(1 ID)
AIC	312.1058
BIC	324.8777
LogLikel.	-152.0529
R-squared Marginal	0.0147
R-squared Conditional	0.5319
Converged	yes
Optimizer	bobyqa

Appendix X Table X.5.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
Time	5.58	1	90.6	0.020

Note. Satterthwaite method for degrees of freedom

Appendix X Table X.5.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.602	0.0562	5.4918	5.712	90.0	99.61	< .001
Time1	4 - 1	0.149	0.0629	0.0253	0.272	90.6	2.36	0.020

Appendix X Table X.5.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.442	0.196	0.525
Residual		0.421	0.177	

Note. Number of Obs: 180, groups: ID 90

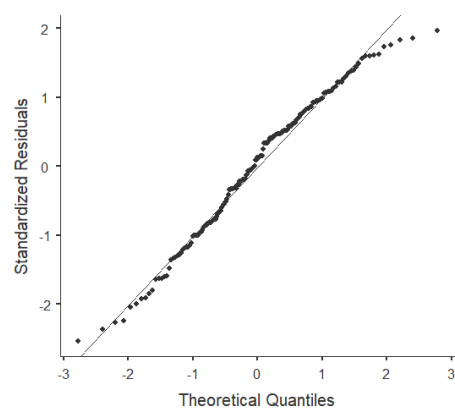
Appendix X Table X.5.5. Post Hoc Comparisons - Time

Comparison		Time	Time	Difference	SE	t	df	p _{bonferroni}
Time								
1	-	4		-0.149	0.0632	-2.35	91.5	0.021

Appendix X Table X.5.6. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0901	0.108

Appendix X Figure X.5.7. Q-Q Plot



Appendix X.6- Impact on Groups

Appendix X Table X.6.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 + Group + Time: ALSA + Time: ALSE+(1 ID)
AIC	218.462
BIC	244.005
LogLikel.	-101.231
R-squared Marginal	0.436
R-squared Conditional	0.746
Converged	yes
Optimizer	bobyqa

Appendix X Table X.6.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
Group	4.39	1	97.4	0.039
Time * ALSA	14.99	2	135.2	< .001
Time * ALSE	14.82	2	137.2	< .001

Note. Satterthwaite method for degrees of freedom

Appendix X Table X.6.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.3067	0.04414	5.22016	5.3932	95.0	120.23	< .001
Group1	G - I	-0.1779	0.08492	-0.34436	-0.0115	97.4	-2.10	0.039
Time1 * ALSA	Time1 * ALSA	0.3019	0.09031	0.12494	0.4789	140.1	3.34	0.001
Time4 * ALSA	Time4 * ALSA	0.4948	0.10508	0.28882	0.7007	139.6	4.71	< .001
Time1 * ALSE	Time1 * ALSE	0.0195	0.00371	0.01222	0.0268	172.0	5.25	< .001
Time4 * ALSE	Time4 * ALSE	0.0138	0.00455	0.00489	0.0227	172.1	3.03	0.003

Appendix X Table X.6.3. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.344	0.1187	0.550
Residual		0.312	0.0973	

Note. Number of Obs: 180 , groups: ID 90

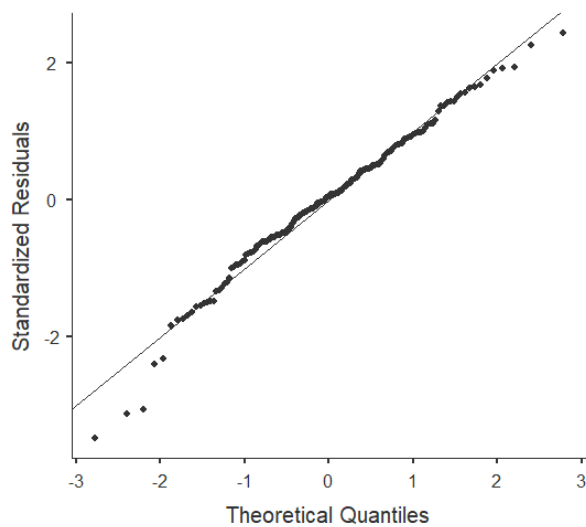
Appendix X Table X.6.3. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	7.00	248	31.4	1.00	< .001

Appendix X Table X.6.3. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0582	0.576

Appendix X Table X.6.3. Q-Q Plot



Appendix Y - KLC's Path Model of Impact

Appendix Y.1- YLE Impact Basic Model

Appendix Y Table Y.1.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 + MRM + ALSR + ALSA + ALSE + Time + MRM: ALSE+(1 ID)
AIC	525.693
BIC	565.174
LogLikel.	-252.847
R-squared Marginal	0.341
R-squared Conditional	0.725
Converged	yes
Optimizer	bobyqa

Appendix Y Table Y.1.2. Fixed Effect Omnibus tests

	F	Num df	Den df	p
MRM	8.19	1	373	0.004
ALSR	25.61	1	379	< .001
ALSA	20.97	1	369	< .001
ALSE	13.24	1	383	< .001
Time	1.83	2	258	0.163
MRM * ALSE	4.48	1	376	0.035

Note. Satterthwaite method for degrees of freedom

Appendix Y Table Y.1.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.22400	0.03808	5.14936	5.2986	164	137.17	< .001
MRM	MRM	0.17552	0.06132	0.05534	0.2957	373	2.86	0.004
ALSR	ALSR	0.21111	0.04171	0.12935	0.2929	379	5.06	< .001
ALSA	ALSA	0.26604	0.05810	0.15217	0.3799	369	4.58	< .001
ALSE	ALSE	0.00818	0.00225	0.00377	0.0126	383	3.64	< .001
Time1	3 - 1	0.06541	0.04529	-0.02335	0.1542	267	1.44	0.150
Time2	4 - 1	0.09771	0.05438	-0.00887	0.2043	274	1.80	0.073
MRM * ALSE	MRM * ALSE	0.00765	0.00361	5.68e-4	0.0147	376	2.12	0.035

Appendix Y Table Y.1.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.409	0.167	0.582
Residual		0.346	0.120	

Note. Number of Obs: 383, groups: ID 160

Appendix Y Table Y.1.5. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	9.00	616	92	1.00	< .001

Appendix Y Table Y.1.5. Random Effect LRT

Test	Statistics	p
Kolmogorov-Smirnov	0.0433	0.471

Appendix Y.2- YLE Impact Intermediate Model

Appendix Y Table Y.2.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 + MRM + Time: ALSA + Time: ALSE + Time: ALSR + MRM: ALSE +(1 ID)
AIC	529.695
BIC	584.967
LogLikel.	-250.847
R-squared Marginal	0.359
R-squared Conditional	0.745
Converged	yes
Optimizer	bobyqa

Appendix Y Table Y.2.2. Fixed Effect Omnibus Tests

	F	Num df	Den df	p
MRM	10.41	1	376	0.001
Time * ALSA	12.73	3	290	< .001
Time * ALSE	4.86	3	303	0.003
Time * ALSR	9.79	3	291	< .001
MRM * ALSE	3.53	1	376	0.061

Note. Satterthwaite method for degrees of freedom

Appendix Y Table Y.2.3. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.20435	0.03847	5.12895	5.2798	161	135.28	< .001
MRM	MRM	0.19787	0.06132	0.07769	0.3181	376	3.23	0.001
Time1 * ALSA	Time1 * ALSA	0.23245	0.06898	0.09725	0.3677	301	3.37	< .001
Time3 * ALSA	Time3 * ALSA	0.41955	0.08234	0.25818	0.5809	308	5.10	< .001
Time4 * ALSA	Time4 * ALSA	0.35515	0.12528	0.10962	0.6007	278	2.83	0.005
Time1 * ALSE	Time1 * ALSE	0.00724	0.00281	0.00174	0.0127	350	2.58	0.010
Time3 * ALSE	Time3 * ALSE	0.00835	0.00327	0.00195	0.0147	335	2.56	0.011
Time4 * ALSE	Time4 * ALSE	0.01117	0.00398	0.00336	0.0190	316	2.80	0.005
Time1 * ALSR	Time1 * ALSR	0.27254	0.05316	0.16835	0.3767	339	5.13	< .001
Time3 * ALSR	Time3 * ALSR	0.14055	0.06041	0.02215	0.2589	324	2.33	0.021
Time4 * ALSR	Time4 * ALSR	0.17287	0.06968	0.03631	0.3094	296	2.48	0.014
MRM * ALSE	MRM * ALSE	0.00683	0.00363	-2.90e-4	0.0139	376	1.88	0.061

Appendix Y Table Y.2.4. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.418	0.175	0.602
Residual		0.340	0.116	

Note. Number of Obs: 391, groups: ID 167

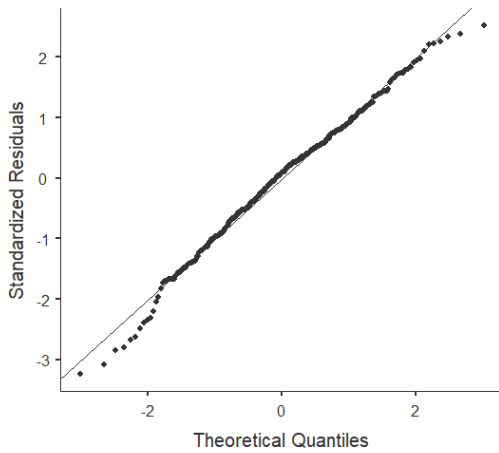
Appendix Y Table Y.2.5. Random Effect LRT

Test	N. par	AIC	LRT	df	p
(1 ID)	13.0	624	95.9	1.00	< .001

Appendix Y Table Y.2.6. Test for Normality of Residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0434	0.468

Appendix Y Table Y.2.7. Q-Q Plot



Appendix Y.3- YLE Impact Advanced Model

Appendix Y Table Y.3.1. Model Info

Info	
Estimate	Linear mixed model fit by ML
Call	ALB ~ 1 + MRM + ALSA +Time: ALSE +Time: ALSR +MRM:ALSE+(1 + MRM:ALSE ID)
AIC	529.325
BIC	592.493
LogLikel.	-248.662
R-squared Marginal	0.360
R-squared Conditional	0.754
Converged	yes
Optimizer	bobyqa

Appendix Y Table Y.3.2. Fixed Effect Omnibus Tests

	F	Num df	Den df	p
MRM	10.02	1	288.1	0.002
Time * ALSA	12.31	3	268.6	< .001
Time * ALSE	4.65	3	230.6	0.004
Time * ALSR	10.98	3	286.4	< .001
MRM * ALSE	4.94	1	39.1	0.032

Note. Satterthwaite method for degrees of freedom

Appendix Y Table Y.3.2. Fixed Effects Parameter Estimates

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	5.20356	0.03928	5.12658	5.2805	152.9	132.48	< .001
MRM	MRM	0.19404	0.06131	0.07387	0.3142	288.1	3.16	0.002
Time1 * ALSA	Time1 * ALSA	0.22263	0.06817	0.08902	0.3562	296.5	3.27	0.001
Time3 * ALSA	Time3 * ALSA	0.41580	0.08131	0.25643	0.5752	303.2	5.11	< .001
Time4 * ALSA	Time4 * ALSA	0.32436	0.12477	0.07982	0.5689	232.6	2.60	0.010
Time1 * ALSE	Time1 * ALSE	0.00723	0.00282	0.00171	0.0128	328.2	2.57	0.011
Time3 * ALSE	Time3 * ALSE	0.00756	0.00331	0.00107	0.0141	273.1	2.28	0.023
Time4 * ALSE	Time4 * ALSE	0.01193	0.00404	0.00401	0.0199	175.8	2.95	0.004
Time1 * ALSR	Time1 * ALSR	0.28357	0.05240	0.18088	0.3863	323.2	5.41	< .001
Time3 * ALSR	Time3 * ALSR	0.16407	0.06023	0.04603	0.2821	259.1	2.72	0.007
Time4 * ALSR	Time4 * ALSR	0.17158	0.06914	0.03608	0.3071	296.2	2.48	0.014
MRM * ALSE	MRM * ALSE	0.00837	0.00377	9.85e-4	0.0157	39.1	2.22	0.032

Appendix Y Table Y.3.3. Random Components

Groups	Name	SD	Variance	ICC
ID	(Intercept)	0.4271	0.182	0.620
	MRM * ALSE	0.0114	1.30e-4	
Residual		0.3343	0.112	

Note. Number of Obs: 383 , groups: ID 160

Appendix Y Table Y.3.4. Random Parameters Correlations

Groups	Param.1	Param.2	Corr.
ID	(Intercept)	MRM	-0.679

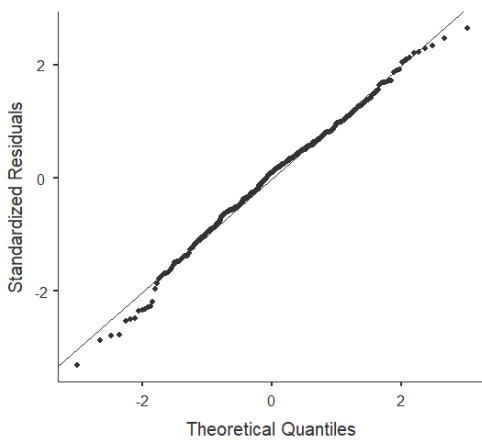
Appendix Y Table Y.3.5. Random Effect LRT

Test	N. par	AI C	LRT	df	p
MRM: ALSE in (1 + MRM: ALSE ID)	14	530	4.37	2.00	0.112

Appendix Y Table Y.3.6. Test for Normality of residuals

Test	Statistics	p
Kolmogorov-Smirnov	0.0478	0.346
Shapiro-Wilk	0.9911	0.020

Appendix Y Table Y.3.1. Q-Q Plot



Appendix Z - Research Protocols and IRB

Appendix Z.1- IRB Approval (Last Modification)



TO: Tim Steffensmeier
School of Leadership Studies
Manhattan, KS 66506

FROM: Lisa Rubin, Chair
Committee on Research Involving Human Subjects

DATE: 01/08/2024

RE: Proposal #IRB-11331, entitled "Validation of Adaptive Leadership Scale."

MODIFICATION OF IRB PROTOCOL #IRB-11331, ENTITLED, "Validation of Adaptive Leadership Scale"

EXPIRATION DATE: Exempt

The Committee on Research Involving Human Subjects (IRB) has reviewed and approved the request identified above as a modification of a previously approved protocol. **Please note that the original expiration remains the same.**

All approved IRB protocols are subject to continuing review at least annually, which may include the examination of records connected with the project. Announced in-progress reviews may also be performed during the course of this approval period by a member of the University Research Compliance Office staff. Unanticipated adverse events involving risk to subjects or to others must be reported immediately to the Chair of the IRB, and / or the URCO

It is important that your human subjects activity is consistent with submissions to funding / contract entities. It is your responsibility to initiate notification procedures to any funding / contract entity of any changes in your activity that affects the use of human subjects.

Electronically signed by Lisa Rubin on 01/08/2024 3:11 PM ET

Appendix Z.2- Consent Forms

Leadership Competencies Questionnaire Informed Consent

(Target Participants: KLC's YLE and LFC participants)

You are invited to participate in KLC's Your Leadership Edge/Lead for Change programs evaluation study, which is led by Ahmadreza Shamsi Yousef, doctoral student in leadership communication at Kansas State University. This study helps to better understand the impact that participation in a leadership development program has on participants' leadership competencies.

You were selected as a participant in this study because you are participant of a leadership development program at Kansas Leadership Center and are at least 18 years of age or older. Your assistance is special because you are developing unique leadership skills and capacities in your respective leadership development program. Your unique perspective and experience will help inform the purpose of this research study.

Participation in this survey will take approximately 15 minutes. Besides this survey, you will be asked to complete three other surveys, which should each take approximately 15 minutes. The link to follow up questionnaires will be distributed one after this program, one 30 days after the program, and one 3-6 months after. Following this letter, you will find an informed consent form, which represents a voluntary agreement to participate in this research project. Upon consent, you will be directed to the first survey. The questions revolve around the way you make decision to practice leadership, your thought processes, beliefs in your leadership capabilities, and your leadership practices. There are no foreseeable risks. While there might be no direct benefit to you, we hope that this research will contribute to designing a scale that broadens our understanding of the impact leadership development programs have on employees. Since the results are expected to show us what factors affect people's ability in experimenting with leadership, you may also indirectly benefit from the result of this research in your organization or community.

All your responses will be anonymous and kept strictly confidential so please be candid and honest. Once we collect the data, each response will receive a unique code and email addresses will be permanently deleted. The data will be stored on a secure password-protected Microsoft platform. The researcher may publish the results of the study, use the data for future studies, and share them with other researchers. If he does, he will benefit from the anonymous data and will only discuss the aggregate results driven from all surveys. In fact, there will never be information that could identify you.

Participation in this study is completely voluntary and you may withdraw at any time without explanation or penalty. While we highly appreciate your contribution to this research, there is no compensation for your participation.

If you have any questions or concerns regarding this study or your rights as a participant, please contact **Ahmadreza Shamsi Yousefi** at Keyhan@ksu.edu or 233 Staley School of Leadership, Kansas State University, Manhattan, KS 66506, (785) 4730602.

You may also contact **Tim Steffensmeier** at steffy@ksu.edu or (785) 532-6099, KLC's director of research and Professor at Kansas State University, Staley School of Leadership.

Also, if you have questions or wish to discuss on any aspect of the research with an official of Kansas State University or the IRB, please contact **Lisa Rubin**, Chair, Committee on Research Involving Human Subjects, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224; **Brad Woods**, Associate Vice President for Research Compliance, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224.

Informed Consent- Adaptive Leadership Scale Content Validity Questionnaire

(Target Participants: Panel of experts)

You are invited to participate in KLC's Your Leadership Edge/Lead for Change programs evaluation study, which is led by Ahmadreza Shamsi Yousef, doctoral student in leadership communication at Kansas State University. The main purpose of this study is to validate and standardize adaptive leadership scales. Moreover, this study helps to better understand the impact that participation in a leadership development program has on participants' leadership competencies.

You were selected as a participant in this survey because you are teaching, studying, or practicing the adaptive approach to leadership and are at least 18 years of age or older. Your assistance is special because you are developing and promoting unique leadership skills and capacities in your work. Your unique perspective and experience will help inform the purpose of this research study.

Participation in this survey will take approximately 30 minutes. Following this letter, you will find an informed consent form, which represents a voluntary agreement to participate in this survey. Upon consent, you will be directed to the survey. In this survey, you will be asked to rate 90 adaptive leadership items in terms of relevancy, clarity, simplicity, and ambiguity. The items are categorized into four main scales: adaptive leadership self-awareness, self-efficacy, reflectiveness, and behavior. There are no foreseeable risks. While there might be no direct benefit to you, we hope that this research will contribute to defining a scale that broadens our understanding of the impact leadership development programs have on people. Since the results are expected to show us what factors affect people's ability in experimenting with leadership, you may also indirectly benefit from the result of this research in your teaching, studies, and work environment.

All your responses will be anonymous and kept strictly confidential. Once we collect the data, each response will receive a unique code and email addresses will be permanently deleted. The data will be stored on a secure password-protected Microsoft platform. The researcher may publish the results of the study, use the data for future studies, and share them with other researchers. If he does, he will benefit from the anonymous data and will only discuss the aggregate results. In fact, there will never be information that could identify you.

Participation in this study is completely voluntary and you may withdraw at any time without explanation or penalty. While we highly appreciate your contribution to this research, there is no compensation for your participation.

If you have any questions or concerns regarding this study or your rights as a participant, please contact **Ahmadreza Shamsi Yousefi** at Keyhan@ksu.edu or 233 Staley School of Leadership, Kansas State University, Manhattan, KS 66506, (785) 4730602.

You may also contact **Tim Steffensmeier** at steffy@ksu.edu or (785) 532-6099, KLC's director of research and Professor at Kansas State University, Staley School of Leadership.

Also, if you have questions or wish to discuss on any aspect of the research with an official of Kansas State University or the IRB, please contact **Lisa Rubin**, Chair, Committee on Research Involving Human Subjects, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224; **Brad Woods**, Associate Vice President for Research Compliance, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224.

Informed Consent- Adaptive Leadership Scale Face Validity Focus Group Discussion

(Target Participants: KLC Teachers)

You have been invited to participate in a focus group discussion designed by Ahmadreza Shamsi Yousefi, doctoral student in leadership communication at Kansas State University. The purpose of this focus group is to test the face validity of Adaptive Leadership scale items identified by the researcher. The information learned in this focus group will be used to finalize the scale items for the evaluation of KLC's Your Leadership Edge/Lead for Change programs

As part of this study, you will be placed in a group of 6 – 12 individuals. A moderator will ask your comments and suggestions on the scale items one by one to make it a valid measure of the adaptive leadership concepts. The discussion will continue until a final consensus on the wordings of the item be reached. This focus group will be recorded on Zoom and a note-taker will be present. However, your responses will remain confidential, and no names will be included in the final report. Please note that there are no right or wrong answers to focus group questions. The researcher wants to hear the many varying viewpoints and would like for everyone to contribute their thoughts. Out of respect, please refrain from interrupting others. However, feel free to be honest even when your responses counter those of other group members. If you choose to participate, you will be asked to respect the privacy of other focus group members by not disclosing any content discussed during the study.

While there might be no direct benefit to you, we hope that this research will contribute to defining a scale that broadens our understanding of the impact leadership development programs have on people. Since the results are expected to show us what factors affect people's ability in experimenting with leadership, you may also indirectly benefit from the result of this research in your teaching, studies, and work environment

Participation in this study is completely voluntary and you may withdraw at any time without explanation or penalty. While we highly appreciate your contribution to this research, there is no compensation for your participation.

If you have any questions or concerns regarding this study or your rights as a participant, please contact **Ahmadreza Shamsi Yousefi** at Keyhan@ksu.edu or 233 Staley School of Leadership, Kansas State University, Manhattan, KS 66506, (785) 4730602.

You may also contact **Tim Steffensmeier** at steffy@ksu.edu or (785) 532-6099, KLC's director of research and Professor at Kansas State University, Staley School of Leadership.

Also, if you have questions or wish to discuss on any aspect of the research with an official of Kansas State University or the IRB, please contact **Lisa Rubin**, Chair, Committee on Research Involving Human Subjects, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224; **Brad Woods**, Associate Vice President for Research Compliance, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224.

Leadership Competencies Questionnaire Informed Consent

(Target Participants: Anybody without leadership background- EFA and CFA Tests)

You are invited to participate in a research study which aims to design a scale for evaluating people's leadership competencies. This research is led by Ahmadreza Shamsi Yousefi, doctoral student in leadership communication at Kansas State University.

You were selected as a participant in this study because you are above 18 and have not participated in any leadership development programs and your insight can help us design a generalizable scale. Your assistance is special because your feedback can help us understand to what extent our scale statements are meaningful.

Participation in this survey will take approximately 15 minutes. Following this letter, you will find an informed consent form, which represents a voluntary agreement to participate in this research project. Upon consent, you will be directed to the survey. The questions revolve around the way you make decision to practice leadership, your thought processes, beliefs in your leadership capabilities, and your leadership practices. There are no foreseeable risks. While there might be no direct benefit to you, we hope that this research will contribute to designing a scale that broadens our understanding of the impact leadership development programs have on people.

All your responses will be anonymous and kept strictly confidential so please be candid and honest. Once we collect the data, each response will receive a unique code and email addresses will be permanently deleted. The data will be stored on a secure password-protected Microsoft platform. The researcher may publish the results of the study, use the data for future studies, and share them with other researchers. If he does, he will benefit from the anonymous data and will only discuss the aggregate results driven from all surveys. In fact, there will never be information that could identify you.

Participation in this study is completely voluntary and you may withdraw at any time without explanation or penalty. While we highly appreciate your contribution to this research, there is no compensation for your participation.

If you have any questions or concerns regarding this study or your rights as a participant, please contact **Ahmadreza Shamsi Yousefi** at Keyhan@ksu.edu or 233 Staley School of Leadership, Kansas State University, Manhattan, KS 66506, (785) 4730602.

You may also contact **Tim Steffensmeier** at steffy@ksu.edu or (785) 532-6099, KLC's director of research and Professor at Kansas State University, Staley School of Leadership.

Also, if you have questions or wish to discuss on any aspect of the research with an official of Kansas State University or the IRB, please contact **Lisa Rubin**, Chair, Committee on Research Involving Human Subjects, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224; **Brad Woods**, Associate Vice President for Research Compliance, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224.

Informed Consent- Labeling Adaptive Leadership Factors Focus Group Discussion

(Target Participants: KLC Teachers and Leadership Faculty/Scholars)

You have been invited to participate in a focus group discussion designed by Ahmadreza Shamsi Yousefi, doctoral student in leadership communication at Kansas State University. The purpose of this focus group is to interpret and label the theoretical factors/constructs identified by the researcher during his work on developing new adaptive leadership scales. The information learned in this focus group will be used to finalize the scale items for the evaluation of KLC's leadership development programs.

As part of this study, you will be placed in a group of 6 – 12 individuals. A moderator will ask you to make sense of the groups of scale items one by one. He will ask you *“when you read the items what theoretical themes did emerge for you related to adaptive leadership framework?”* The discussion will continue until a final consensus on the themes be reached. This focus group will be recorded hybrid (Zoom and in-person at the Staley School of Leadership) and a note-taker will be present. However, your responses will remain confidential, and no names will be included in the final report. Please note that there are no right or wrong answers to focus group questions. The researcher wants to hear the many varying viewpoints and would like for everyone to contribute their thoughts. Out of respect, please refrain from interrupting others. However, feel free to be honest even when your responses counter those of other group members. If you choose to participate, you will be asked to respect the privacy of other focus group members by not disclosing any content discussed during the study.

While there might be no direct benefit to you, we hope that this research will contribute to defining a scale that broadens our understanding of the impact leadership development programs have on people. Since the results are expected to show us, what factors affect people's ability in experimenting with leadership, you may also indirectly benefit from the result of this research in your teaching, studies, and work environment

Participation in this study is completely voluntary and you may withdraw at any time without explanation or penalty. While we highly appreciate your contribution to this research, there is no compensation for your participation.

If you have any questions or concerns regarding this study or your rights as a participant, please contact **Ahmadreza Shamsi Yousefi** at Keyhan@ksu.edu or 233 Staley School of Leadership, Kansas State University, Manhattan, KS 66506, (785) 4730602.

You may also contact **Tim Steffensmeier** at steffy@ksu.edu or (785) 532-6099, KLC's director of research and Professor at Kansas State University, Staley School of Leadership.

Also, if you have questions or wish to discuss on any aspect of the research with an official of Kansas State University or the IRB, please contact **Lisa Rubin**, Chair, Committee on Research Involving Human Subjects, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224; **Brad Woods**, Associate Vice President for Research Compliance, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224.

Informed Consent- Qualitative Interviews

(Target Participants: KLC/LFC Participants)

You are invited to complete your participation in the KLC's leadership development path 2023 study. The purpose of this study is to better understand the impact that participation in a leadership development program has on participants' leadership behaviors. This study will provide the foundation for a better understanding about leadership in organizational contexts as well as communities.

You were selected as a participant in this study because you are a past participant of a leadership development program in the state of Kansas and are at least 18 years of age or older. Your assistance is special because you have developed unique leadership skills and capacities in your respective leadership development program. Your unique perspective and experience will help inform the purpose of this research study.

Participation in this survey and interview will take approximately 60-90 minutes in maximum two sessions. The questions revolve around your stories of success, failure, and non-action in experimenting with leadership since completing the YLE program of Kansas Leadership Center (KLC). All these three kinds of stories are important and valuable to us. The interviewer will ask you about your specific experiences, thoughts, and feelings while practicing leadership in your work environment or community. There are no foreseeable risks. While there might be no direct benefit to you, we hope that this research will contribute to our understanding of the impact leadership development programs have on employees and help us improve them. Since the results are expected to show us what factors affect people's ability in experimenting with leadership, you may also indirectly benefit from the result of this research in your organization or community.

All your responses will be kept strictly confidential so please be candid and honest. This information will only be seen by the researchers. Interview recording and its transcription will be stored on a secure Microsoft platform. The researchers may publish the results of the study. If they do, they will only discuss the aggregate results driven from all interviews. Moreover, before publishing, the participants will receive a zero draft of a research summary to voluntarily read and give their final views. Your name will not be used in any publication or presentation about the study. We may share the data we collect from you for use in future research studies or with other researchers – if we share the data that we collect about you, we will remove any information that could identify you before we share it.

Participation in this study is completely voluntary and you may withdraw at any time without explanation or penalty. While we highly appreciate your contribution to this research, there is no compensation for your participation.

If you have any questions or concerns regarding this study or your rights as a participant, please contact **Ahmadreza Shamsi Yousefi** at Keyhan@ksu.edu or 233 Staley School of Leadership, Kansas State University, Manhattan, KS 66506, (785) 4730602.

You may also contact **Tim Steffensmeier** at teffy@ksu.edu or (785) 532-6099, KLC's director of research and Professor at Kansas State University, Staley School of Leadership.

Also, if you have questions or wish to discuss on any aspect of the research with an official of Kansas State University or the IRB, please contact **Lisa Rubin**, Chair, Committee on Research Involving Human Subjects, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224; **Brad Woods**, Associate Vice President for Research Compliance, 203 Fairchild Hall, Kansas State University, Manhattan, KS 66506, (785) 532-3224.