Cultivating moral reasoning:

An empirical study on the impact of an ethical dilemma reflection framework

Maureen Spelman

Cultivating moral reasoning:

An empirical study on the impact of an ethical dilemma reflection framework.

Maureen Spelman

Professor, Educational Leadership
North Central College

Abstract

Leaders engaging in the processes of making decisions, judgments, and/or navigating ethical dilemmas are often challenged by elements of ambiguity or even conflicting values. Learning to handle challenging situations takes time and practice. The use of ethical dilemma case studies in school leadership programs can provide the opportunity to expose leadership candidates to real world dilemmas, authentic decision-making, and enable them to size up problems while considering the broader school community. Such embedded experiences have the potential to guide aspiring leaders as they apply theory in practice, cultivate the capacity for critical analysis and decision-making, along with building adaptive expertise, muscle memory, and overall ethical fitness. This study explores the initial phase of an unfolding curriculum continuum focused on the impact of using an ethical dilemma reflection framework to cultivate moral reasoning in aspiring school leaders. Data was gathered from the Defining Issues Test (DIT2) using a pre-, post, post-post experimental design. Quantitative data demonstrates that the DIT2 N2 scores show a trajectory of growth with some instances of statistical significance in the demonstration of moral reasoning. This classroom-based study of practice and effectiveness of intervention may be of interest and relevance to wider audiences.

Universities have a long history of training leaders, but in the unique paradigm of the 21st century it has become critically apparent that leaders, particularly in education, need to be intentionally prepared to lead with a moral compass. Traditionally, educational leadership programs in the U.S. have focused on the competencies needed to lead effectively in complex and challenging environments. If these programs are to produce leaders who are not only competent managers, but also wise thinkers and ethical leaders, then institutions of higher education need to have an equal emphasis on developing the virtue of practical wisdom in future leaders who will guide schools in this increasingly global context (Sturm, Vera, and Crossan, 2017; Brooks, Brant, and Lamb, 2019; Jubilee Centre & Oxford Character Project, 2020; Lamb, Brandt, and Brooks, 2021).

Practical Wisdom

Phronesis or practical wisdom is a foundational virtue ethics concept that refers to choosing the right course of action. What Aristotle called phronesis is key to the process of making the right choices. He took the stance that practical wisdom is necessary for every virtue to function with a moral compass and all other virtues are united under the umbrella of this meta-virtue (Aristotle, 1999). Essentially, demonstrating good character demands that the individual practices well-informed judgment in considering the relative weight of competing virtues (Arthur et al., 2021). Schwartz and Sharpe (2010) agree, and argue that practical wisdom is what prevents virtues from becoming vices.

Wright, Warren, and Snow (2021) state that any discussion of virtue must consider the place of practical wisdom in the development and integration of virtues into the character of an individual. They argue that the exercise of virtue is informed, guided, and mediated by practical wisdom. Kristjánsson et al., (2021) describes practical wisdom as a meta-virtue that supports the capacity of knowing and choosing the correct moral action. While some theorists suggest that virtues are developed in childhood and adolescence, Sanderse (2020) maintains that the intentional cultivation of virtues can continue throughout adulthood.

Leaders engaging in the processes of making decisions, judgments, and/or navigating ethical dilemmas are often challenged by elements of ambiguity or even conflicting values. School leaders at all levels frequently struggle and search for the correct path to guide their decisions in complex circumstances. Schwartz and Sharpe (2010) suggest that practical wisdom is a key skill in the process of making the right choices. Hackett and Wang (2012) and Newstead et al. (2019) agree that developing the virtue of practical wisdom should be a significant consideration in the design of leadership preparation programs.

According to Spillane and Lowenhaupt (2019) recognizing and managing dilemmas is central to the work of school leaders; and knowledge of the rules alone will not be sufficient when school leaders are faced with the need to deliberate and adjudicate in determining the proper course of action. They note that school leaders regularly face complicated dilemmas that derive from a clash of values, and are expected to craft solutions that are grounded in wise moral choices (Spillane & Lowenhaupt 2019). Learning to handle challenging situations takes time and practice. A study of codes, standards, and policies will not sufficiently guide leaders to make wise choices. Institutions of

higher education, then, need to prepare leadership candidates for more than a specific job or role; there is a critical need to engage in the character education and virtue development of aspiring leadership candidates (Jubilee Centre and Oxford Character Project 2020; Lamb, Brandt, and Brooks, 2021).

Collective Wisdom

> *For it is possible that the many, though not individually good men,*
>
> *yet when they come together may be better, not individually but*
>
> *collectively, than those who are so, just as public dinners to which*
>
> *many contribute are better than those supplied at one man's cost;*
>
> *for where there are many, each individual, it may be argued, has*
>
> *some portion of virtue and [phronesis]*
>
> — Aristotle, *Politics* (1999)

In *Nicomachean Ethics,* Aristotle (1999) discusses *phronesis* in the context of an individual grappling with moral issues. However, in *Politics* (1944) he provides an extended discussion of *phronesis* in the context of rulers dealing with competing virtues at the state level. Kristjánsson et al., (2021) point out that city states in Aristotle's time were similar in size to modern medium-sized institutions. Thus, Aristotle's comments regarding the "many" may have implications for our modern U.S. schools that embrace the format of leadership teams and distributed leadership. Aristotle observes that the "many" in this sense are less corruptible and less likely to fall prey to negative passions or vices (Kristjánsson et al., 2021).

Gary and Judith Olsen have spent decades studying the ways in which individuals think together. They suggest that group cognition is enhanced through conversations combined with tangible representations of the task to be completed (2000). Olsen and Olsen's research revealed the benefits of having group thinking made visible for all participants allowing for reflection, and revision in community. In his book, *The Wisdom of Crowds,* James Surowiecki (2004) notes that groups of individuals often come to better, collectively wiser decisions than individuals or those who might be

considered the smartest. Surowiecki (2004) explains notion of the wisdom of crowds as the idea that large groups of people are collectively smarter than individual experts when it comes to problem-solving and decision-making. His research findings underscore the point that the more important the decision, the greater the need to tap into the perspectives of a team of leaders.

As Carl Wiemann (2007) searched for a way to support his physics students to be more dexterous thinkers he stumbled upon the value of social engagement. His investigation found that development of crucial skills was closely related to students' engagement with peers (2007). HWiemann's work suggests that intelligent decision-making is a social process. More recently, the concepts of distributed cognition and collective intelligence have been explored by Annie Murphy Paul (2021). In her recent book, *The Extended Mind,* Murphy Paul notes that technologies now enable scientists to scan individuals' brains as they interact. As a result, they have found fascinating evidence to support the premise that when we interact socially our brains engage different neural and cognitive processes than when we think or act on our own (Murphy Paul, 2021). Her analysis of thinking with peers suggests that we think differently, and often better in groups.

The concept of collective *phronesis* may have significant implications for programs preparing leaders for the current climate of joint decision-making. Kristjánsson et al. (2021) emphasizes that programs that prepare leaders need to do more than teach about phronesis: they need to provide aspiring leaders opportunities to engage in and practice collective phronesis through joint decision-making exercises.

Reflection and Dialogue

Scholars agree that through reflection and collegial dialogue, adult learners share learning and begin to develop common understandings and thus effectively deconstruct assumptions. Brookfield (1995, 2000) referenced critical reflection as a primary goal of adult education. Jarvis (1987, 2001), later expanded upon this discussion and defined reflective learning as the practice of planning, monitoring, and reflecting upon experiences.

Daloz (1999), Freire (2000) and Mezirow (2000) agree that through collegial dialogue, adult learners share learning and begin to develop common understandings. Johnson-Bailey and Alfred (2006) also

explored the value of dialogue and found that adult learners effectively deconstruct assumptions through dialogue with others. Current theorists Drago-Severson and Blum-DeStefano (2018) suggest that collegial conversations add an informed layer to the reflective process. From the literature, it is apparent that these themes are critical elements to consider when preparing school leaders.

Context

Opportunities to practice and reflect in community can empower educators to navigate critical incidents with a moral compass. Practicing with and responding to authentic ethical dilemmas through a structured reflective framework may be one strategy to help educators engage in collective phronesis to make sense of complexity, and to choose the intelligent and virtuous response to ethical dilemmas. In this study educational leadership candidates engaged in exercises related to ethical decision making to develop virtue reasoning. Intervention participants broke into small groups to engage in authentic education ethical dilemmas. Candidates followed a four-phase ethical dilemma reflection framework (Appendix A) as they debated, discussed, and attempted to come to consensus on solutions. Note-takers recorded conversation highlights as the group engaged in dialogue following the four-phase reflection framework. The full class then reconvened to compare, discuss, and debate the responses generated by each small group to agree on the best course of action.

Method

Research Design

This research study tests the hypothesis that groups exploring education-based dilemmas using an ethical dilemma reflection framework will demonstrate advancements in the development of moral reasoning in the initial course of an educational leadership program. A secondary exploration tests the hypothesis that candidates will maintain and/or continue to improve moral reasoning scores at the end of the two-year program in which the use of ethical dilemma reflection framework is practiced and reinforced. This study uses a pre- and post-test design for the data gathered in the initial course and a pre- post-post design at the end of the full two-year program. Data were collected across three years and five discrete cohorts of educational leadership candidates (2021-2023)

The hypotheses derive from my personal viewpoint as a subjective constructionist with an interpretivist approach aiming to understand the impact of character-led professional judgment (Cohen, Manion, & Morrison, 2011). Brown-Urban, et al., (2014) suggest four distinct phases that describe a progression of program evolution. The level of program maturity for this participatory action research study is situated in the initiation phase of the evolutionary continuum (Brown-Urban, et al., 2014). The initiation phase label describes programs that are either new or existing ones that have been significantly revised in a new context. Accordingly, the recommended phase one evaluations would fall under the category of process and response. Process and response evaluations tend to be dynamic, flexible, and can provide timely feedback to guide program adjustments (Brown-Urban, et al., 2014). Additionally, the lens of viable validity allows the researcher to explore the sustainability of program implementation beyond the timeframe and resources of the research study.

Participants

As is often seen in the education field, it was not possible to gather a random group of participants. Candidates taking part in this ongoing research study include a convenience sampling composed of 48 participants from five different cohorts enrolled in an educational leadership master's program. All study participants completed the DIT2 pre-assessment prior to starting program coursework and again at the conclusion of the first course. In addition, 15 candidates from two cohorts have completed the post-post DIT2 assessment at the conclusion of the capstone course. The remaining candidates will complete the post-post DIT2 in coming semesters as they complete the educational leadership program. A sixth cohort of 35 candidates have taken the pre-assessment and will complete the post-assessment at the end of the fall 2023 academic term.

Table 1. Demographic statistics of study participants.

	N	Male/Female
Cohort 3 – summer 2021	7	2/5
Cohort 4 – fall 2021	8	2/6
Cohort 5 – spring 2022	7	2/5

| Cohort 6 – fall 2022 | 19 | 8/11 |
| Cohort 7 – spring 2023 | 7 | 1/6 |

Ethical Considerations

Ethical approval for this study was granted by the Institutional Review Board at North Central College. The researcher's intent to address trustworthiness was intentionally embedded in each component of this study. Participation in this study was strictly voluntary; there were no penalties if candidates chose not to participate. The researcher took precautions to ensure the privacy and confidentiality of study participants. To ensure that information was collected in an ethical manner, each piece of data was assigned an identification number.

Measure

The data collected for this study includes a pre/post and pre/post-post comparison of the Defining Issues Test (DIT2). The DIT has been in use since the 1970's, and according to Rest, et al., (2000) the validity of this instrument is based on seven specific criteria and hundreds of studies that have produced significant trend results. The DIT2 is a revised version of the original DIT first developed by Rest, Cooper, Coder, Masanz, and Anderson (1974). Essentially, the DIT and DIT2 consist of scenarios concerning social problems.

The DIT2 assessment consists of rating and ranking tasks designed to assess how an individual interprets and responds to moral dilemmas. As they engaged in the DIT2, participants in this study were presented with five moral dilemma scenarios and decide what choice the protagonist should make. Participants then rated a set of 12 statements in terms of importance on a 5-point scale. Finally, they identified the four statements that best described and supported their choice of how the protagonist should act. The assumption in this assessment of moral reasoning is that individuals define the most important issue of a dilemma according to the level of their moral reasoning (Cummings, et al., 2010). The presumption is that moral judgment consists of three schemas that are developmentally ordered: personal interest, maintaining norms, and post conventional (Rest, et al., 2000).

The principled reasoning score (P Score) represents an individual's relative location on the continuum and is focused on the post conventional schema items. The newest index, N2, can be described as a modified P Score. The N2 index uses the P Score as a starting point, and then adjusts that score based on the individual's ability to discriminate between that post conventional score and items lower on the developmental continuum (Dong & Thoma, 2014). In this study, the participants' pre- DIT2 N2 scores served as a baseline indicator of moral reasoning. While the DIT2 moral dilemmas are not education specific, DIT researchers suggest that the exercise of rating and ranking of statements for all five scenarios can provide an index of how the participant may approach moral decision making in their professional context (Dong & Thoma, 2014; Roche & Thoma, 2017). In addition, Dong & Thoma (2014) submit that the DIT2 measure is sensitive to educational intervention.

Process

Each participant completed the pre- and post- DIT2 assessments online via Qualtrics links. The initial analysis of the DIT2 data was conducted by the Center for the Study of Ethical Development at the University of Alabama. The pre- and post-reports for both leadership candidates include developmental indices, developmental profile and phase indices, experimental indices, means, and standard deviations for personal interest (stage two and three), maintain norms (stage four), P Scores, and N2 Scores. The N2 score reports were then analyzed using a t-test to determine significance difference between the pre- and post- scores. This type of inferential statistic is often used in hypothesis testing to determine whether a process or treatment has influenced the population of interest. If a p-value reported from a t-test is less than 0.05, then that result is said to be statistically significant.

In addition, tests of significance were supplemented with a measure of effect size. The decision to explore the effect size allows the researcher to determine practical significance independent of the sample size. Cohen's *d* is a known test statistic that quantifies mean differences in standard deviation units. Cohen (1969) describes an effect size of 0.2 as small, 0.5 as medium, and 0.8 as large. In this study, Cohen's *d* was calculated for the combined cohorts 3-7.

Findings
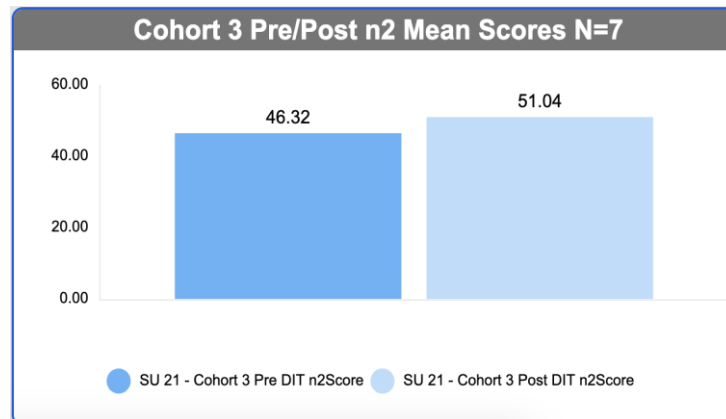
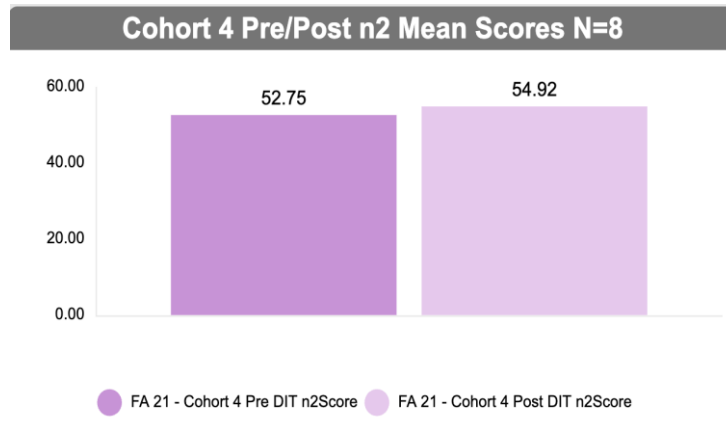|		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre n2	46.3171	7	11.70879	4.42551
	Post n2	51.0400	7	6.41017	2.42282

		Upper	t	df	Sig. (2–tailed)
Pair 1	Pre n2 – Post n2	.82726	−2.082	6	.082

The 2-tailed t-test results demonstrate that there is no significance between the pre- and post-test on the DIT2 performance for the 7 summer 2021 candidates. For this group the null hypothesis is retained, however, there is a positive trajectory in the performance on the DIT2 over the course of the term.

Cohort 4 – Fall 2021

Cohort 4 Pre/Post n2 Mean Scores N=8

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre n2	52.7500	8	10.57963	3.74046
	Post n2	54.9188	8	9.49019	3.35529

		Upper	t	df	Sig. (2-tailed)
Pair 1	Pre n2 – Post n2	2.39866	−1.123	7	.299

The fall 2021 data once again failed to demonstrate significance between the pre- and post-test DIT2 N2 score of the 8 candidates. For this group the null hypothesis is retained. It is interesting to note, however, that this cohort's pre-test results were high in comparison to the other four. That being noted, the trajectory for this cohort still moves in a positive direction.
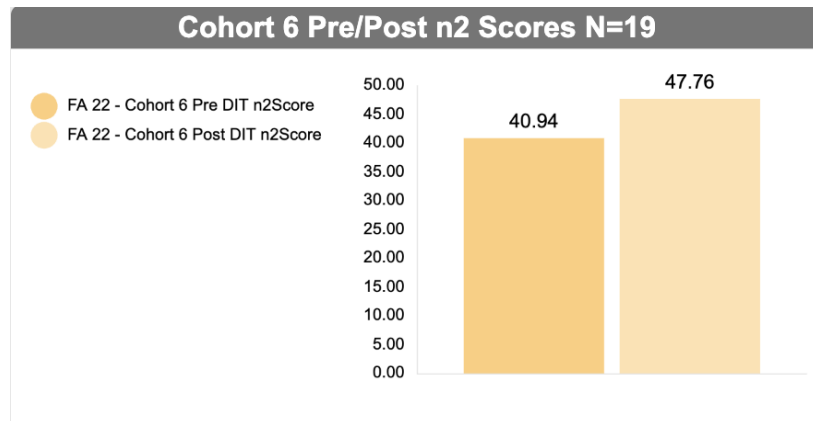
Cohort 5 – Spring 2022

Cohort 5 Pre/Post n2 Scores N=7
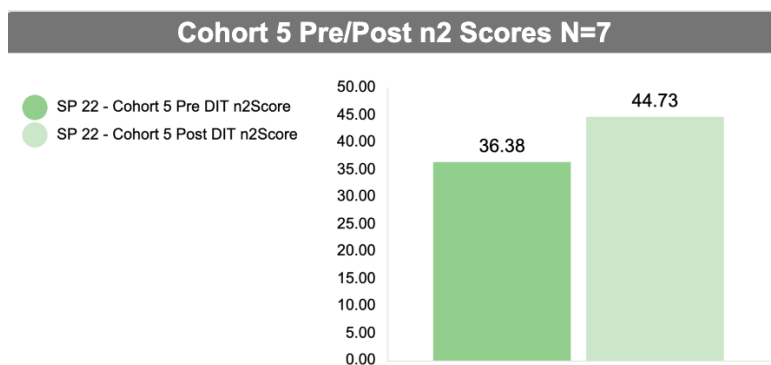
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre n2	36.3786	7	9.10437	3.44113
	Post n2	44.7286	7	7.52021	2.84237

		Upper	t	df	Sig. (2-tailed)
Pair 1	Pre n2 – Post n2	-.37986	-2.564	6	.043

The 2-tailed t-test results demonstrate for the spring 2022 cohort demonstrate significance and thus reject the null hypothesis in favor of the experimental hypothesis. The results suggest that the ethical dilemma interventions had an impact on the spring 2022 cohort candidates' performance on the DIT2 post assessment.

Cohort 6 – Fall 2022

Cohort 6 Pre/Post n2 Scores N=19

FA 22 - Cohort 6 Pre DIT n2Score
FA 22 - Cohort 6 Post DIT n2Score

40.94   47.76

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre n2	40.9405	19	13.23294	3.03585
	Post n2	47.7553	19	9.34247	2.14331

		Upper	t	df	Sig. (2-tailed)
Pair 1	Pre n2 – Post n2	−2.19396	−3.098	18	.006

The 2-tailed t-test results demonstrate for the fall 2022 cohort also demonstrate significance and thus reject the null hypothesis in favor of the experimental hypothesis. The results suggest that the ethical dilemma interventions had an impact on the fall 2022 cohort candidates' performance on the DIT2 post assessment.
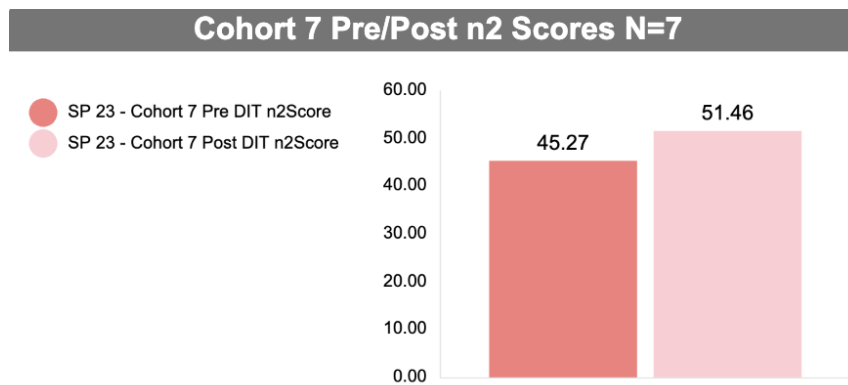
Cohort 7 – Spring 2023

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre n2	45.2700	7	14.83022	5.60530
	Post n2	51.4571	7	15.62537	5.90584

		Upper	t	df	Sig. (2-tailed)
Pair 1	Pre n2 – Post n2	3.55733	−1.554	6	.171

For the spring 2023 cohort the 2-tailed t-test results demonstrate that there is no significance between the pre- and post-test on the DIT2 performance for the 7 summer 2021 candidates. For this group the null hypothesis is retained, however, there is a strong positive trajectory in the performance on the DIT2 between the pre- and post-test administrations.

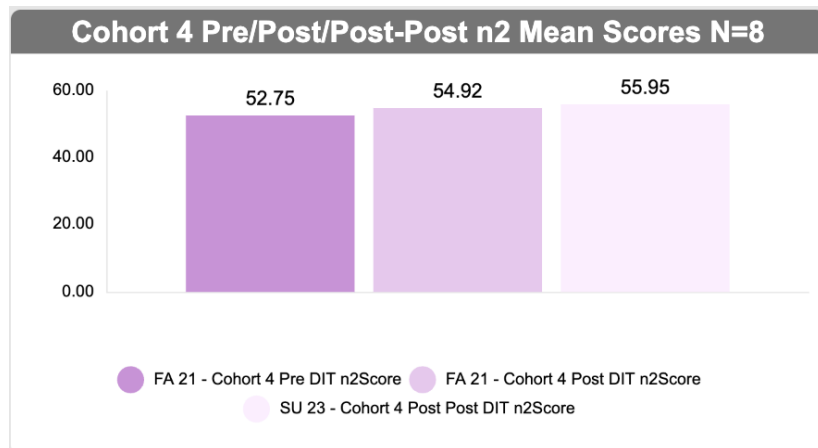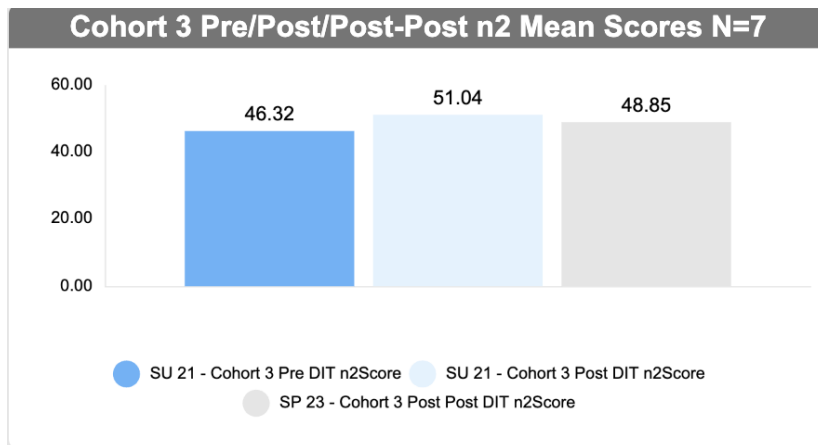*Combined Cohorts 3, 4, 5, 6, & 7 – pre- and post*

2021-2023 Cohorts Pre/Post n2 Scores N=48
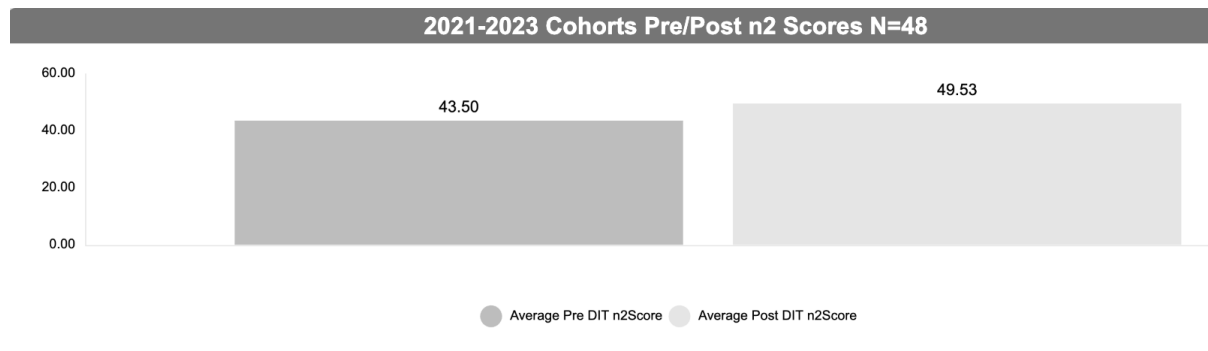
		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre n2	43.6590	48	12.89504	1.86124
	Post n2	49.5267	48	10.05070	1.45069

		N	Correlation	Sig.
Pair 1	Pre n2 & Post n2	48	.755	.000

The 2-tailed t-test results for the combined cohorts demonstrate significance and thus reject the null hypothesis in favor of the experimental hypothesis. The overall positive trajectory suggests that the ethical dilemma interventions had an impact on cohort candidates' performance on the DIT2 post assessment. To determine the distributional separation for the combined cohorts the Cohen's *d was* calculated and yielded a *d* of -.69 (a medium effect). The r-square of 33% indicates that 33% of the variation can be explained by the intervention implemented between the pre- and post-test scores.

Cohorts 3 and 4 – pre-, post, and post-post

Cohort 3 Pre/Post/Post-Post n2 Mean Scores N=7

46.32   51.04   48.85

SU 21 - Cohort 3 Pre DIT n2Score   SU 21 - Cohort 3 Post DIT n2Score
SP 23 - Cohort 3 Post Post DIT n2Score

Cohort 4 Pre/Post/Post-Post n2 Mean Scores N=8

52.75   54.92   55.95

FA 21 - Cohort 4 Pre DIT n2Score   FA 21 - Cohort 4 Post DIT n2Score
SU 23 - Cohort 4 Post Post DIT n2Score

Cohorts 3 & 4 Pre/Post-Post n2 Scores N=15

49.75   53.11

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pre n2	49.7480	15	11.21404	2.89545
	PostPost n2	52.6393	15	13.59322	3.50975

		Upper	t	df	Sig. (2-tailed)
Pair 1	Pre n2 – PostPost n2	2.91492	−1.068	14	.304

Fifteen candidates from cohorts 3 and 4 have completed the educational leadership program and the post-post DIT2 assessment at the conclusion of the capstone course. For the combined cohorts, the 2-tailed t-test results demonstrate that there is no significance between the pre- and post-post-test on the DIT2 performance for the 15 candidates at the conclusion of the program. For these groups the null hypothesis is retained, however, there is a positive trajectory in the performance on the DIT2 between the pre- and post-post-test administrations.

Discussion

There is some evidence in support of the initial hypothesis that groups of candidates exploring education-based dilemmas using an ethical dilemma reflection framework will demonstrate advancements in the development of moral reasoning. The evidence appears to support the initial hypothesis across the five semesters of combined cohort data. Cohorts 5 and 6 showed statistically significant improvement between pre- and post-test DIT2 scores, while cohorts 3, 4, and 7 demonstrated increased scores and positive trajectory.

In terms of the secondary hypothesis that candidates will maintain and/or continue to improve moral reasoning scores at the end of the two-year program in which the use of ethical dilemma reflection framework is practiced and reinforced the data is mixed. The post-post DIT2 scores for cohort 3 declined from the post-test scores. The cohort 4 scores continued in a positive trajectory from the pre- and post-test scores. However, the overall findings provide empirical support to the value of ethical dilemma activities designed to develop the moral reasoning and collective phronesis in future school leaders.

The results suggest several questions for discussion. First, what might explain the fact that cohorts 5 and 6 showed substantial quantitative improvement in their post DIT2 scores? What might be the reason behind the positive trajectory, but less substantial improvement in the pre- and post-test scores for cohorts 3, 4, and 7? What factors might explain the decline in post-post scores for cohort 3? Finally, what future considerations might help to improve results for cohorts still in the pipeline as well as future cohorts?

Fidelity of Implementation

Implementation fidelity is the degree to which an intervention is delivered as intended and is critical to successful translation of evidence-based interventions into practice. Berkowitz and Bier (2007) discuss the critical importance of ensuring fidelity when implementing character and virtue-based interventions. In this study the initial course in the educational leadership program is taught by the same instructor for each cohort of candidates. This initial course, entitled Mission Driven Leadership, was redesigned with a focus on virtue ethics and piloted in summer 2021. It might be safe to posit that the implementation of the revised curriculum strengthened as the semesters progressed, thus possibly explaining the increasing trajectory of DIT2 post-test scores over time.

Another factor to consider might be the lived experiences of candidates choosing to enroll in this character and virtue-based leadership program. Candidates enrolled in the five different cohorts come from a wide variety of school contexts and roles. Some candidates come to this leadership program because they are attracted to the strong emphasis on ethical leadership, while others enroll due to scholarships offered to educators working in diverse schools and/or from underrepresented populations. Essentially, candidates enter this program with a wide range of personal and professional experience that may impact their moral reasoning scores on the pre-DIT2 assessment. Each cohort represents a unique set of aspiring school leaders.

It is important to investigate possible reasons for the fact that cohort 3 post-post DIT2 scores dipped slightly from the post-test scores. Cohort 3 began the educational leadership program in the summer of 2021 as a group of seven new candidates. However, as they moved through the educational leadership program, they were joined at various points by five additional candidates who were 'off-track' and had started the program prior to the curriculum revisions. The five additional candidates did not have the same foundational knowledge offered in the revised initial course and may not

have been on the same page in classroom conversations and ethical dilemma activities. Also, cohort 3 experienced one course with an instructor that briefly joined the program faculty for just one semester. Another new faculty member was hired the following semester, and this newest faculty member was in the process of orienting to the character and virtue-based curriculum while teaching courses for cohort 3 candidates. Another senior faculty member was awarded a sabbatical and was replaced by adjunct instructors for two of the courses for cohort 3. These combined circumstances may have resulted in the Cohort 3 candidates experiencing inconsistencies in the implementation of the Ethical Dilemma Reflection Framework as they moved through the program.

Implementation fidelity is critical to successful translation of evidence-based interventions into practice.

A lack of implementation fidelity in the case of cohort 3 may have been a factor in the intervention being less effective for this group over the course the full program. If the intervention was not delivered as conceived and planned it may explain the decline in the average post-post test scores. Thus, the results suggest that future work might focus on regular training and professional development for program instructors. In addition, implementation support and frequent communication with program faculty may help to ensure the fidelity of implementation going forward. Ensuring regular practice in reflective engagement with authentic ethical dilemmas in community may help candidates to develop an ethical fitness (Kidder, 2009) that will transfer to their future professional roles.

Limitations

There are several limitations that must be considered in this research study. To begin, this study was not intended to present generalizable findings, and the conclusions are applicable only to the unique program redesigned for this educational leadership program. That being noted, the findings may also be of interest to other institutions and programs focused on the development of moral leaders. While character theories suggest that the development of moral reasoning is life-long work, most support the belief that growth is more modest at the adult level than in childhood or adolescence even when intentional interventions are implemented (Grant, 2010; Thompson, 2015; McAdams, 2015; Carr, 2017).

The study is also limited due to the small data set and the fact that the convenience sampling presents the possibility of bias. This small sample size does not provide the quantity of data needed to extrapolate findings to larger populations. That being noted, this early stage study of the first three years of implementation may be joined with future cohorts to increase its contribution to discourse on the development of moral reasoning in aspiring leaders.

This study was undertaken from the inside by one of the redesign architects who also serves as the instructor for the initial course, thus, the potential for bias also exists in this paradigm. The researcher's personal stance as a subjective constructivist combined with an interpretivist approach certainly could influence how study results have been interpreted. The aim was to conduct a research study that will have direct relevance to professional practice in delivery of this newly redesigned educational leadership program (Munn-Giddings, 2017; Punch & Oancea, 2014; Robson & McCartan, 2016; Sagor, 2000). Ultimately, the data was filtered through a personal set of theoretical frames and biases.

It is important to note that this research study primarily explored the impact of character-based education interventions in the first term of a two-year educational leadership program. As such, it represents a snapshot in time. Given the developmental nature of virtue, candidates' experience in the initial course merely represents the beginning of a journey (Annas, 2011). The data on cohorts completing this two-year program is still being collected and analyzed.

In addition to guiding future program delivery, this study also raises the need for ongoing, longitudinal research. The intent is to conduct follow-up studies with candidates in years one, three, and five of their professional leadership roles to investigate whether gains have been maintained. The measure utilized for the proposed longitudinal study will be an Intermediate Concept Measure recently developed at the University of Alabama. This new measure will incorporate dilemmas relevant to specific educational leadership settings.

Conclusion

This manuscript has analyzed the qualitative data representing one effort to cultivate moral reasoning in an educational leadership program. The findings of this small study have the potential

to spark a larger conversation around the value of intentionally infusing leadership preparation programs with opportunities to grapple with and reflect on ethical dilemmas in community. Recent studies have explored how character and virtues inform professional practice, but very little can be found in the literature that specifically relates to the development moral reasoning in school leaders. Professional educators work in constantly changing environments that require careful discernment as well as operating withing professional codes of conduct, rules, and regulations.

Through a virtue ethics lens, moral reasoning requires an awareness of relevant and competing virtues. Practical wisdom or phronesis is thought to be the essence of good judgment and involves evaluating the relative value of competing virtues, identifying perspectives and emotions of all stakeholders, and remaining open to various options for action. Ethical leaders are sensitive to the context and the stakeholders in the community they serve.

Aristotle wrote about the need to focus on virtue development, and phronesis in particular, as a path to living and flourishing in community. In a similar vein, Linda Zagzebski (1996, 2004), a contemporary virtue theorist, observed that virtuous individuals can have greater impact than those focused solely on rules and regulations. As Aristotle and Zagzebski suggest, those who lead with integrity and embody virtues such as practical wisdom are more likely to assemble a loyal cadre of followers. Along the same line, scholars Sergiovanni (1984) and Starratt (2004) believed that virtue ethics can serve as a valuable foundation for the development of education leaders who serve with a moral compass. Wagner and Simpson (2009) added to the literature supporting the need for both competence and character as they argued that the heart of educational leadership begins with the moral decision-making and impacts the overall moral architecture of the school community.

Professional educators that lead with a moral compass need to carefully consider the context, critically evaluate the relative weight of competing virtues, examine the perspectives of all stakeholders, and remain open-minded to potential solutions. A valuable method for developing moral decision-making processes may be through sustained practice with authentic ethical dilemmas. Practice in groups using an Ethical Dilemma Reflection Framework may support the development of moral reasoning and produce leaders who have internalized a moral compass, value the benefits of collective wisdom and be better poised to implement character-led professional judgment in future leadership roles.

Acknowledgements

Appendix A

Ethical Dilemma Reflection Framework (Spelman, 2023)

Phase I – Be Present (name and strive to understand the situation & perspectives applicable to the situation – ethic of critique)
- Share initial reactions.
- Formulate questions.
- What professional codes of ethics, school rules and/or policies need to be consulted?
- What virtues are relevant to the dilemma?
- Are different virtues in conflict, and if so, which should be prioritized in this situation? Why?
- List all possible stakeholders.

Phase II – Frame & Reframe (keep the focus of the process on the people and their best interests – ethic of care)
- ✓ Examine the emotions, interests, relationships, and perspectives of all stakeholders.
- ✓ Explore the values, virtues, and/or beliefs that may impact various stakeholder perceptions.
- ✓ Identify and focus on what's essential, ignore the noise,
- ✓ Recognize and name biases.
- ✓ Determine player(s) that own the moral issue (s).

Phase III – Engage in Dialogue (deciding on actions that will maximize benefits for all while also respecting the rights of individuals – ethic of justice)

- ✓ What decision would the profession expect?
- ✓ What decision would the community expect?
- ✓ What is the right decision based on what is best for students?
- ✓ Identify Option A and Option B
- ✓ Brainstorm creative alternatives, is there an Option C?
- ✓ Come to consensus and defend the final decision.

Phase IV – Reflect on the Process

- ✓ Does your final decision differ from your initial reactions?
- ✓ How did your personal and professional ethics influence the final decision?
- ✓ What leadership style(s) would an ethical and moral leader need to call upon to deal with this dilemma?
- ✓ What leadership virtues and/or character strengths an ethical and moral leader need to call upon to deal with this dilemma?

References

Annas, J. (2011) *Intelligent virtue.* Oxford: Oxford University Press.

Aristotle. (1944). *Politics.* (H. Rackham, W. Heinemann, Trans). Harvard University Press.

Aristotle. (1999). *Nicomachean ethics.* (2nd ed. T. Irwin. Trans). Hackett Publishing.

Arthur, J., Earl, S., Thompson, A., & Ward, J. (2021). The value of character-based judgement in

the professional domain. *Journal of Business Ethics. 169*, 293-308.

https://doi.org/10.1007/s10551-019-04269-7

Berkowitz, M. & Bier, M., (2007). What works in character education. *Journal of Research in Character*

    Education, 5(1), 29-48.

Brookfield, S. (1995). *Becoming a critically reflective teacher.* Jossey-Bass.

Brookfield, S. (2000). Transformative learning as ideology critique. In Mezirow, J. (Ed.)

    Learning as transformation: Critical perspectives on a theory in progress. (pp. 125-148).

    Jossey-Bass.

Brooks, E., Brant, J., and Lamb, M. (2019). How can universities cultivate leaders of character?

    Insights from a leadership and character development program at the University of Oxford.

    Journal of Ethics Education, 4(2), 167-182.

Brown-Urban, J., Hargraves, M., & Trochim, M. (2014). Evolutionary evaluation: Implications

    for evaluators, researchers, practitioners, funders, and the evidence-based program

    mandate. *Evaluation and Program Planning, 45,*127-139.

Carr, D. (2017). Virtue and character in higher education. *British Journal of Educational Studies, 65*(1),

    109-124.

Cohen, J. (1969). Statistical Power Analysis for the Behavioral Sciences (1st ed.). New York: Academic

    Press.

Cohen, L., Manion, L., & Morrison, K. (2011). *Qualitative inquiry and research design:*

    Choosing among five approaches (3rd ed.) Sage Publications.

Cummings, R., Maddux, C., Cladianos, A., & Richmond, A. (2010). Moral reasoning of

    Education students: The effects of direct instruction in moral development theory and

    participation in moral dilemma discussion. *Teachers College Record, 112*(3), 621–644.

Daloz, L. (1999). *Mentor: Guiding the journey of adult learners* (2nd ed.) Jossey-Bass.

Dong, Y. & Thoma, S. (2014). The defining issues test of moral judgment development. *APA Behavioral Development Bulletin, 19*(3), 55-61.

Drago-Severson, E. & Blum-DeSefano, J. (2018). *Leading change together: Developing educator capacity within schools and systems.* Association for Supervision and Curriculum Development.

Freire, P. (2000). *Pedagogy of the oppressed* (20th anniversary edition). Continuum.

Grant, R. (2010). Is humanistic education humanizing? In E. Kiss & J. Euben (Eds.), *Debating moral education: Rethinking the role of the modern university.* (pp. 283-295). Duke University Press.

Hackett, R. & Wang, G. (2012). Virtues and leadership: An integrating conceptual framework founded in Aristotelian and Confucian perspectives on virtues. *Management Decision*, *50*(5), 868-899.

Jarvis, P. (1987). *Adult learning in the social context.* Croom Helm.

Jarvis, P. (2001). *Learning in later life: An introduction for educators and careers.* Kogan Page.

Johnson-Bailey, J., & Alfred, M. (2006). Transformational teaching and the practice of Black women adult educators. in Taylor, E. (Ed.), *Fostering transformative learning in the classroom: Challenges and innovations. New directions in adult and continuing education*, (pp. 49-58) Jossey-Bass.

Jubilee Centre for Character and Virtues and Oxford Character Project. (2020). Character education in universities. A framework for flourishing. *Jubilee Centre for Character and Virtues. University of Birmingham,* 1-12. https://oxfordcharacter.org/uploads/files/Character-Education-in-Universities.pdf

Kidder, R. (2009) *How good people make tough choices: Resolving the dilemmas of ethical living.* New York: Harper.

Kristjánsson, K., Fowers, B., Darnell, C., & Pollard, D. (2021). Phronesis (practical Wisdom) as a type of contextual integrative thinking. *Review of General Psychology, 25*(3), 239-257. https://doi.org/10.1177/10892680211023063

Lamb, M., Brandt, J., Brooks, E. (2021). How is virtue cultivated? Seven strategies for postgraduate character development. *Journal of Character Education, 17*(1), 81-108.

McAdams, D. (2015). Psychological science and the nicomachean ethics: Virtuous actors, agents, and authors. In N. Snow (Ed.), *Cultivating virtue: Perspectives from philosophy, theology, and psychology.* Pp. 307-336). Oxford University Press.

Mezirow, J. (2000). Learning to think like an adult: Core concepts of transformation theory. In Mezirow, J. (Ed.) *Learning as transformation: Critical perspectives on a theory in progress* (pp. 3-33). Jossey-Bass.

Munn-Giddings, C. (2017). Action research, In R. Coe, M. Waring, L. Hedges, & J. Arthur (Eds.), *Research Methods & Methodologies in Education (*pp. 71-77). Sage Publications.

Murphy-Paul, A. (2021). *The extended mind: The power of thinking outside the brain.* Mariner.

Newstead, T., Dawkins, S., Macklin, R., & Martin, A. (2019). We don't need more leaders – We need more *good* leaders. Advancing a virtues-based approach to leader(ship) development. *The Leadership Quarterly, 32*, 1-11. http://doi.org/10.1016/j.leaqua.2019.101312

Olsen, G. & Olsen, J. (2000). Distance matters. Human computer interaction. 15, 139-178.

Punch, K., & Oancea, A. (2014) *Introduction to research methods in education. 2nd edition.* Sage.

Rest, J., Cooper, D., Coder, R., Masanz, J. & Anderson, D. (1974). Judging the important moral
   issues in moral dilemma – an objective test of development. *Developmental Psychology,
   10*(4), 491-501.

Rest, J., Narvaez, D., Thoma, S., & Bebeau, M., (2000). A neo-Kohlbergian approach to
   morality research. *Journal of Moral Education, 29*(4), 381-395.

Robson, C. & McCartan, K. (2016) *Real world research: A resource for users of social research
   methods in applied settings 4th edition.* John Wiley & Sons Ltd.

Roche, C. & Thoma, S. (2017). Insights from the Defining Issues Test on moral reasoning
   Competencies development in community pharmacists. *American Journal of Pharmaceutical
   Education, 81*(8), 5913 DOI: 10.5688/ajpe5913.

Sagor, R. (2000) *Guiding school improvement with action research.* Association for Supervision
   and Curriculum Development.

Sanderse, W. (2020). Does Aristotle believe that habituation is only for children? *Journal of
   Moral Education, 49(*1), 98-110.

Schwartz, B. & Sharpe, K. (2010). *Practical wisdom: The right way to do the right thing.*
   Riverhead Books.

Sergiovanni, T. (1984). Leadership and excellence in schooling. *Educational Leadership, 41*
   (5). (4-13).

Spelman, M. (June 2023). *Ethical dilemma reflection framework.* Jubilee Centre for Character &
   Virtues – Teaching character: A practical guide. Available at:
   https://www.jubileecentre.ac.uk/2983/character-education/framework/framework-in-action

Spillane, J. & Lowenhaupt, R., (2019). *Navigating the principalship: Key insights for new and aspiring school leaders.* Association for Supervision and Curriculum Development.

Starratt, R. (2004) *Ethical leadership.* Jossey-Bass.

Sturm, R., Vera, D., & Crossan, M. (2017). The entanglement of leader character and leader competence and its impact on performance. *The Leadership Quarterly, 28*, 349-366.

Surowiecki, J. (2004). *The wisdom of crowds.* Anchor Books.

Wagner, P. & Simpson, D. (2009). *Ethical decision making in school administration: Leadership as moral architecture.* Sage.

Wiemann, C. (2007). Why not try a scientific approach to science education? *Change: The Magazine of Higher Learning. 39,* 9-15.

Wright, J., Warren, M., & Snow, N. (2021). *Understanding virtue: Theory and measurement.* Oxford University Press.

Zagzebski, L. (1996) *Virtues of the mind.* Cambridge University Press.

Zagzebski, L. (2004) *Divine motivation theory.* Oxford University Press.