CULTIVATING VIRTUE IN POSTGRADUATES: AN EMPIRICAL STUDY OF THE OXFORD GLOBAL LEADERSHIP INITIATIVE

Submission for *The Journal of Moral Education*

Corresponding Author:

Jonathan Brant, DPhil
University of Oxford
jonathan.brant@theology.ox.ac.uk; +44 7854 771041
Faculty of Theology & Religion, Gibson Building, Woodstock Rd, OX2 6GG.

Jonathan David Brant (DPhil) is Founding Director of The Oxford Character Project, an interdisciplinary initiative that combines academic research on character development with a practical programme for University of Oxford postgraduate students. He is also a Research Fellow of Harris Manchester College and a member of the Faculty of Theology and Religion at the University of Oxford.

Co-Authors:

Michael Lamb, PhD
Wake Forest University
lambkm@wfu.edu; +1 336 758 2343
P. O. Box 7225, Winston-Salem, NC 27109, USA

Michael Lamb (PhD) is Assistant Professor of Politics, Ethics, and Interdisciplinary Humanities and Director of the Program for Leadership and Character at Wake Forest University. He is also a Research Fellow for the Oxford Character Project. His research focuses on virtue ethics, character education, and the ethics of citizenship.

Emily Burdett, DPhil
University of Nottingham
emily.burdett@nottingham.ac.uk; +44 1158 467346
School of Psychology, University Park, Nottingham, NG7 2RD

Emily Burdett (DPhil) is Assistant Professor, Faculty of Science, University of Nottingham and also a research affiliate at the Institute of Cognitive and Evolutionary Anthropology, University of Oxford. She pursues research in cognitive science, social cognition, and developmental psychology.

Edward Brooks, DPhil
Oxford Character Project
ed.brooks@oxfordcharacter.org; +44 7967 157068

Edward Brooks (DPhil) is the Executive Director of the Oxford Character Project and also heads the Global Leadership Initiative. His academic work focuses on how the virtue of hope might be understood and cultivated in late modernity.

Abstract: Although virtue ethics has emerged as an influential ethical theory within the academy, universities have not generally taken up the practical task of virtue cultivation. Some academics even resist the effort altogether. In response, this article presents an early-stage evaluation of one effort to cultivate virtue in postgraduate students, a theoretically derived and empirically measured character development programme at the University of Oxford. The study uses a pre- and post-test experimental design to assess whether participation results in measurable growth of four character virtues. Quantitative data offer evidence for modest improvement with respect to two of the four focal virtues measured, while qualitative data suggest that future iterations should ensure that participants are given both reflective, conceptual tools and practical, everyday tasks to cultivate virtue. The article provides both empirical support and future directions for ongoing research and programme development for cultivating virtue in the university.

Keywords: character; virtue; university; higher education; postgraduates; measurement

Acknowledgements: This project was made possible through the support of a grant from Templeton World Charity Foundation Inc. [TWCF0061]. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Templeton World Charity Foundation Inc. We are also grateful to the participants of the Global Leadership Initiative who made this research possible and pleasurable, and, especially, to the staff: Anjie Anderson, Kate Seagrave, Claire Shuttleworth, Julianne Viola, Luna Wang, and Bethan Willis. We appreciate the Trustees of the Oxford Pastorate who shared wisdom and provided institutional support. We are also grateful to Nigel Biggar and the McDonald Centre for Theology, Ethics and Public Life, and to Wake Forest University for supporting this research. We wish to thank the Jubilee Centre for Character and Virtues who organised the conference and seminar where this research was first presented. Eranda Jayawickreme and Laura Blackie supplied early advice on the project, Christian Miller offered constructive feedback, and Cameron Silverglate provided invaluable editorial assistance.

Disclosure statement: The authors have no financial interest or benefit arising from the direct applications of this research.

CULTIVATING VIRTUE IN POSTGRADUATES: AN EMPIRICAL STUDY OF THE OXFORD GLOBAL LEADERSHIP INITIATIVE

In the latter half of the twentieth century, seminal theoretical works across a range of disciplines repositioned virtue ethics as an influential ethical theory within the academy (e.g. Anscombe, 1958; Hauerwas, 1994; MacIntyre, 1981). While virtue ethics was challenged by situationists who doubted the existence of global traits of character (Doris, 2002; Harman, 2000), the energetic riposte from psychologists and philosophers contributed to more nuanced accounts of virtue and character buttressed by the results of empirical investigation (e.g., Fleeson, 2001; Miller, 2013, 2014; Sabini and Silver, 2005; Snow, 2009; Sreenivsan, 2002). Recently, academics have applied theoretical resources to the practical formation of character (e.g., Annas, Narvaez, & Snow, 2016; Peterson & Seligman, 2004; Snow, 2015) and developed instruments to measure specific character virtues (e.g., Morgan, Gulliford, & Kristjánsson, 2017). These developments have informed moral education, encouraging new methods to cultivate and measure character (Kristjánsson, 2015). Ironically, however, universities—the original site of the recovery of virtue ethics—have been reluctant to participate in the intentional task of virtue development. To the extent that universities engage moral education at all, most are content to study character rather than develop programmes to cultivate it.

In some quarters, there has even been explicit resistance to moral education in the university. Most influentially, Stanley Fish (2003a, 2003b, 2004) has pressed a number of criticisms against efforts to form character in the academy. Casting the university primarily as a 'place for teaching and research,' he argues that universities may cultivate 'intellectual virtues' in the pursuit of 'truth' but not moral virtues, which, he argues, are 'tangential' to a university's core educational mission (2003a). According to Fish, 'it is not the business of the university' to encourage moral behaviour since doing so would require the university to determine which moral view is 'the right one' and therefore 'would deform (by replacing) the true task of academic work: the search for truth and the dissemination of it through teaching' (2004). Elsewhere, Fish suggests that moral education encourages a 'discipleship that is itself suspect and dangerous' and argues that forming character is not only a 'bad idea' but an 'unworkable' one: 'There are just too many intervening variables, too many uncontrolled factors that mediate the relationship between what goes on in a classroom or even in a succession of classrooms and the shape of what is finally a life' (2003b). David Carr echoes Fish's concerns in relation to postgraduate character education, wondering whether it is either possible or legitimate for those beyond the age of majority pursuing specialized study or professional training (2017).

Defenders of moral education have offered powerful rebuttals. Elizabeth Kiss and J. Peter Euben have addressed each of Fish's worries, suggesting, for example, that while universities should be modest about their ability to influence students' moral character, examples affirm the transformative role that faculty can have on students' lives (2010a, pp.

60-61). Furthermore, moral education need not be imposed as an indoctrinating form of 'discipleship' but instead can engage 'deliberative and dialogical practices' that encourage students to be more critical and reflective in forming their own judgments (2010a, p. 62; cf. 2010b, pp. 20-21). Moreover, Fish's 'sharp dichotomy' between intellectual and moral virtues ignores the ways that academic integrity often involves broader moral concerns, and his emphasis on producing and sharing academic knowledge neglects the 'vast and varied terrain of general undergraduate education, professional and vocational education, residential life, and extracurricular activities,' all of which shape character without necessarily impinging on the academic mission of the classroom (2010a, pp. 67-68). Given that universities inevitably shape the character of students, Kiss and Euben conclude, the question is not whether they will but 'how, when, and by whom' (2010b, pp. 14, 17; cf. Colby 2002).

This article seeks to bolster support for moral education in the university by presenting an early-stage evaluation of one effort to cultivate character within a university context, the Oxford Global Leadership Initiative (GLI) run by the Oxford Character Project (OCP), a new, theoretically derived and rigorously measured character development programme for postgraduate students at the University of Oxford. By focusing on emerging adults—a population largely neglected in character research focused mostly on children and adolescents—the article seeks to contribute a new perspective to how and when character education might be possible. By analysing a voluntary, extracurricular programme for postgraduates, it seeks to show how such an intervention can be responsive and even immune to the concerns raised by critics of moral education in the university.

The Oxford Global Leadership Initiative

The Oxford Global Leadership Initiative (GLI) seeks to help talented students develop key virtues of character that will prepare them to be the wise thinkers and good leaders that the world needs. Focusing on virtues and practices that are essential for character-based leadership, the GLI utilises a research-based practical programme that draws together cohorts of students to develop qualities of life and leadership in diverse and intellectually open learning communities.

Each GLI learning community consists of 12–14 postgraduate students who have voluntarily applied for the programme and who follow a seven-month curriculum that includes engagement with prepared sets of readings; group seminars, retreats, and discussions; dinner conversations with expert speakers; one-on-one mentoring from programme staff and a senior leader in the student's vocational field; and workshops addressing the arts of leadership, including virtues and practices presented through jazz, Shakespearian drama, and portraiture (see Blinded for Review, 2019a). The voluntary nature of the programme insulates it against concerns about the normative imposition of character on those above the age of majority (Carr, 2017), and its extracurricular nature means it does not divert resources from the academic research and teaching that Fish and others consider to be the only legitimate concern of universities (Fish, 2003a, 2003b, 2004).

The programme components are the practical instantiation of a specific theoretical approach and methodology for character virtue development that underlies the work of the GLI. The model is broadly Aristotelian, referring to good habits of character as virtues and

bad habits as vices and assuming that individual virtues and overall character may be intentionally developed (Aristotle, 1999; Burnyeat, 1980, pp. 86–8; Miller, 2013; Kristjánsson, 2014). In particular, GLI targets four focal virtues—sense of vocation; commitment to service; humility; and gratitude. These virtues were selected because they are not naturally cultivated through the rigours of postgraduate study (as, for example, diligence); they are directed toward common goods, not merely individual self-interest; and they are relevant to the life-stage known as 'emerging adulthood' (Arnett, 2014a, 2014b; for detailed explanation, see Blinded for Review, 2019a). Moreover, these virtues are valued by a wide range of traditions, and participants are encouraged to critically engage the GLI material in light of their personal commitments and their tradition's account of the good, thereby mitigating any danger of 'discipleship' (Fish, 2003b).

The focal virtues were selected, and working definitions were decided, on the basis of the practical experience of the senior academics and leaders involved in establishing the OCP. Nonetheless, theoretical accounts of these virtues have continued to inform the discussion and working definitions outlined in Table 2 below. For example, a sense of vocation is informed by Duffy & Sedlacek (2007), Dik & Duffy (2009), and Dik, Eldridge, Steger & Duffy (2012); a commitment to service by Rushton et al. (1981), Batson (2011, especially pp. 12-32), and Carlo et al (2009); humility by Exline & Hill (2012), Tangney, (2000, 2009), and Hill and Sandage (2016); and gratitude by Emmons & McCullough (2004) and Morgan, Gulliford, & Kristjánsson (2017).

Drawing upon a broad range of research across education, philosophy, psychology, neuroscience, and theology, the GLI utilises seven 'pedagogical strategies' (Berkowitz & Bier, 2007, pp. 39–43, especially 43), which, together, are intended to help postgraduate students cultivate good character and grow towards becoming wise thinkers and good leaders: '1) habituation through practice; 2) reflection on personal experience; 3) engagement with virtuous exemplars; 4) dialogue that increases virtue literacy; 5) awareness of situational variables; 6) moral reminders; and 7) friendships of mutual accountability' (Blinded for Review, 2019a).

The GLI is also committed to evaluative measurement through the use of controlled studies, which can help to address Fish's contention that character education is 'unworkable' and too difficult to assess (2003b). To date, examples of this kind of measurement project remain scarce, but within moral education and the sub-discipline of character education in particular, the literature is growing (Bulach, 2002; Harrington, Giles, Hoyle, Feeney, & Yungbluth, 2001; Lamb & Randazzo, 2016; Leming, 2000; Wang, Ferris, Hershberg, & Lerner, 2015). Both Was, Woltz and Drew (2006) and Berkowitz and Bier (2007) have assessed multiple examples of character programme measurement. While Was et al. (2006) are pessimistic, citing 'conceptual and methodological weaknesses' (p. 148), Berkowitz and Bier (2007) believe that some studies are sufficiently reliable to make it 'clear that character education can effectively promote the development of a wide array of psychological outcomes that can be construed as aspects of character' (p. 41; see also Berkowitz, 2011).

These two meta-analyses enable us to position this measurement project in relation to the scholarly literature and note two ways it will make a contribution to the field. First, with respect to the target population, the GLI is unique in focusing on emerging adults, specifically postgraduate students. Recent studies have shown this to be a time of moral

formation and value determination, situating 18–29-year-olds in a 'transformative period of self-development' (Noftle, 2015, p. 490; cf. Arnett, 2014a, 2014b). Second, this study brings together a number of the desiderata laid out in the meta-analyses, notably the use of psychometrically validated assessment instruments and a longitudinal, mixed-method, quasi-experimental design (see especially Berkowitz & Bier, 2007, p. 44).

Method

Research design

The measurement project tests the hypothesis: those who have invested significant time and energy in a programme intended to develop character (and these four focal virtues in particular) will have modestly improved their scores over the period of the longitudinal studies more than members of a control group who, while sharing a similar motivation, do not have the opportunity to participate. Drawing upon recent discussion of methodological issues in the measurement and evaluation of character education programmes (e.g. Card, 2017 and Deutsch, 2017), the study uses a pre- and post-test, quasi-experimental, longitudinal research design with quantitative data supplemented by qualitative data intended to illuminate the quantitative results.

Participants

Postgraduate students wishing to participate in the GLI completed an online application and submitted a *curriculum vitae*. Approximately sixty-five applications were received for the 13–14 places available each year. Selection was based on a clear understanding of the aims of the programme (especially the way its focus on character differs from skills- or technique-based leadership programmes), evidence of leadership experience, and expressed commitment to attend all events and engage constructively. Given the focus on global leadership in a pluralistic context, there was also an effort to create a diverse group with respect to gender, nationality, academic discipline, and religious and philosophical commitment. The control groups were selected by the same methods and from the same communities and networks, including some who had applied to the programme but had not been included.

Across three years of activity, data was collected from 41 GLI participants (the intervention group) and 23 members of a non-equivalent group drawn from the same intact group (Oxford University postgraduates) who had not engaged with GLI programming (the control group).

Table 1. Demographic statistics of study participants

	N	**Male / Female**	**Nationalities**

GLI Cohort 2015	13	5 / 8	Australia; China; France; Germany; India; Nepal; New Zealand; Singapore; South Africa; USA
GLI Cohort 2016	14	6 / 8	Australia; Canada; China; Columbia; Kenya; Pakistan; South Africa; UK; USA
GLI Cohort 2017	14	8 / 6	Australia; China; Germany; India; Paraguay; UK; USA; Zambia
Control 2015	4	3 / 1	Singapore; UK; USA
Control 2016	10	5 / 5	Canada; India; Pakistan; UK; USA
Control 2017	9	5 / 4	Australia; China; Indonesia; Italy; USA

Procedure

Each year, a questionnaire was administered before and after each postgraduate progressed through the cohort programme. To accord with best practice in human research and encourage candour among participants, questionnaires were completed anonymously. As this was a longitudinal study, each postgraduate was asked to use a pseudonym, which enabled the matching of pre- and post-test responses. The instruments utilised had been checked for consistency over time using test-retest reliability measures (DeVellis, 2012, pp. 51–3). The questionnaire was produced and delivered as a pencil-and-paper test in year 1 and then transferred online for ease of completion and recording of results in years 2 and 3. The possible impact of a change of administration method is discussed in DeVellis (2012, p. 52), but our results show no differences across the two methods. This process was repeated over three years to assess the impact of the GLI programme on each cohort.

To supplement quantitative data from the scale questionnaires, qualitative data was collected from a long-answer questionnaire presented to participants at the end of the programme alongside the post-test scale questionnaires. This qualitative questionnaire included questions about the components and impact of the programme, plus each of the four focal virtues. The questionnaires resembled the schedule for a semi-structured interview with direct questions about the programme followed by prompts and space for reflective, open-ended answers. This mixed-method approach allows the illuminating power of the qualitative data to inform analysis of the quantitative results.

The response rate to questionnaires was high among programme participants, especially in years 2 and 3. Overall, the number of complete sets of responses (pre-test questionnaire, post-test questionnaire, long-answer / qualitative questionnaire) was 32 out of

41, leading to a 78% response rate. In the results and analysis below, we include these 32 responses when examining differences between Time 1 and Time 2.

Control group participants completed the questionnaires at the same pre- and post-test times as the participants. They were incentivised by the offer of a £10 Amazon voucher upon the second (i.e. final) completion of the survey. We aimed to recruit half as many control group members as intervention participants in each year as resources were limited and computer modelling experiments have shown a smaller control group size does not greatly impact on the measurement of intervention effect (Hutchins, Brown, Mayberry, & Sollecito, 2015, especially p. 234). Complete sets of responses (pre- and post-test surveys) were received from 4 control group members in year 1, 7 in year 2, and 6 in year 3, leading to a control N=17, which achieved the desired 50% ratio to intervention participants.

In addition to providing data for this study, the results, especially from the qualitative questionnaire, were used to refine and improve the practical programme. All research was approved by the Central University Research Ethics Committee at the University of Oxford and undertaken in accordance with best practice as recommended by this body with respect to informed consent, data-storage, and complaint procedures.

Measures

We followed the established practice of using proven research instruments validated by extensive psychometric testing of large samples of respondents (DeVellis, 2012, p. 185). We chose four instruments (published or used with permission) with compatible question formats and the same response scale so that they could be combined into a single questionnaire: Brief Calling Scale by Dik, Eldridge, Steger and Duffy (2012); Self-Report Altruism Scale by Rushton, Chrisjohn and Fekken (1991); Humility Scale by Elliott (2010); and Gratitude Scale by McCullough, Emmons and Tsang (2002).

Card (2017) offers a detailed discussion of the measurement of character constructs such as strengths and virtues. Noting that such constructs frequently have 'fuzzy boundaries,' he offers a figure which helps researchers visualise whether there is sufficient congruity between their definition of the character construct and the questionnaire items used to measure the presence of that construct (Card 2017, p. 32, cf. p. 35). Given Card's approach, we believe there is sufficient conceptual and semantic similarity between the GLI's working definitions and the chosen scales' definitions to ensure the individual items and the variables measured by the scales appropriate proxies for the GLIs focal virtues (DeVellis, 2012, p. 60).

Table 2. Comparison of measured traits

From the GLI project description:	**From the instruments we utilise:**

‘Sense of vocation’ – Characteristic of the person who believes herself to have an orienting purpose that transcends mere personal success or flourishing. In secular contexts, vocation may be understood to be the result of a ‘call’ from a particular community. In religious contexts, vocation may be understood to arise out of relationship with God or a divine being(s) who calls individuals to make use of their unique gifts.	‘Calling’ – ‘A person’s belief that she or he is called upon (by the needs of society, one’s own inner potential, by God, by a Higher Power, etc.) to do a particular kind of work.’ (Dik et al., 2012, pp. 253–4)
‘Commitment to service’ – Characteristic of the person who is appropriately other-focused rather than merely self-focused and intends a positive impact and contribution to the common good within her or his wider social and communal context.	‘Altruism’ – ‘[T]here is a trait of altruism. That is, some people are consistently more generous, helping and kind than others.’ ‘[We assessed] four global ratings of the target person's altruism—that is, how caring, how helpful, how considerate of others' feelings and how willing to make a sacrifice the individual was.’ (Rushton et al., 1981, p. 296)
‘Humility’ – Characteristic of the person who is not deceived by pride and sees himself/herself as he/she truly is. This trait makes it possible to see the worth, merit, and value of others and of others’ opinions and beliefs. The humble person will consider others’ needs and be open to new developments and ideas and willing to revise their own positions.	‘Humility’ – ‘An accurate assessment of one’s abilities and achievements; the ability to acknowledge one’s mistakes, imperfections, gaps in knowledge, and limitations; an openness to new ideas, contradictory information, and advice; a keeping of one’s abilities and accomplishments—one’s place in the world—in perspective; a relatively low self-focus, a ‘forgetting of the self’, while recognising that one is but one part of the larger universe; an appreciation of the many different ways that people and things can contribute to our world.’ (Elliott, following Tangney’s definition, 2010, p. 6)

'Gratitude' – Characteristic of the person who is appropriately grateful or thankful and recognises that he or she is not responsible for all the good that they have enjoyed but recognises that others—ancestors, parents, teachers, or peers, for example—have contributed to their life, success, and happiness.	'Gratitude' – 'We define the grateful disposition as a generalised tendency to recognise and respond with grateful emotion to the roles of other people's benevolence in the positive experiences and outcomes that one obtains.' (McCullough et al., 2002, p. 112)

Results

Quantitative Results

Aggregate scores were created for each participant for each virtue. These scores were then averaged so that each virtue could be compared. We first examined whether there were differences in responses according to each virtue among the cohorts for the GLI groups and separately for the control groups. We ran four one-way ANOVA's (for each virtue) with each year of the GLI groups (3: 2015, 2016, 2017) as the between variable and one of the virtues as the within variable. We ran another four one-way ANOVA's with all control groups by each year (3: 2015, 2016, 2017) as the between variable and each virtue as the within variable. There were no significant differences across all three years among the GLI groups or among the control groups, *p*s > .06. Since there were no significant differences, all cohorts within the GLI groups were combined to form one GLI group, and similarly, all control groups were combined to form one control group to increase the statistical power of results.

As a summary of what follows, we first examined the average individual participant scores at pre- (Time 1) and post-test (Time 2) for each virtue by the combined GLI group and then the control group separately. Finally, we compared both the control and GLI groups at Time 1 and Time 2 to evaluate the GLI programme.

Differences between pre- and post-testing by each of the four virtues

GLI Cohort Groups

We grouped items by virtue (Service, Gratitude, Humility, and Vocation) and created an average score for each person by virtue. We conducted paired t-tests to compare pre- and post-test scores for each virtue. We found significant differences between Time 1 and Time 2 for two virtues: Service and Gratitude (see Table 3). Participants in the GLI groups reported significantly higher experiences of Gratitude and Service at post-testing compared to pre-testing.

Table 3. Average scores at Time 1 and 2 for each virtue for the GLI groups combined

	Time 1 *M (SD)*	Time 2 *M (SD)*	*t* score	*p* value	Cohen's *d* (Effect size *r*)
Virtue					
Service	4.01 (.436)	4.26 (.489)	3.75	**.001**	1.35 (.56)
Gratitude	4.45 (.504)	4.73 (.351)	4.33	< **.0001**	1.56 (.61)
Humility	3.09 (.316)	3.19 (.329)	1.53	.135	.55 (.26)
Vocation	3.55 (.697)	3.64 (.745)	.731	.47	.26 (.13)

Control Groups

We also compared scores at Time 1 and Time 2 for the control group for each virtue. Paired t-tests revealed no significant differences for any virtue, *p*s > .244 (see Table 4). For the control groups, there were no changes in reported experiences for any virtue at Time 1 and Time 2.

Table 4. Average scores at Time 1 and 2 for each virtue for the control groups combined

	Time 1 *M (SD)*	Time 2 *M (SD)*	*t* score	*p* value	Cohen's *d* (Effect size *r*)
Virtue					
Service	3.94 (.503)	3.89 (.508)	.552	.589	.28 (.14)
Gratitude	4.31 (.737)	4.39 (.612)	.855	.41	.43 (.21)
Humility	3.02 (.355)	3.04 (.293)	.239	.814	.12 (.06)
Vocation	3.15 (.884)	3.31 (.758)	1.21	.24	.61 (.29)

Comparison of control and GLI groups for each virtue

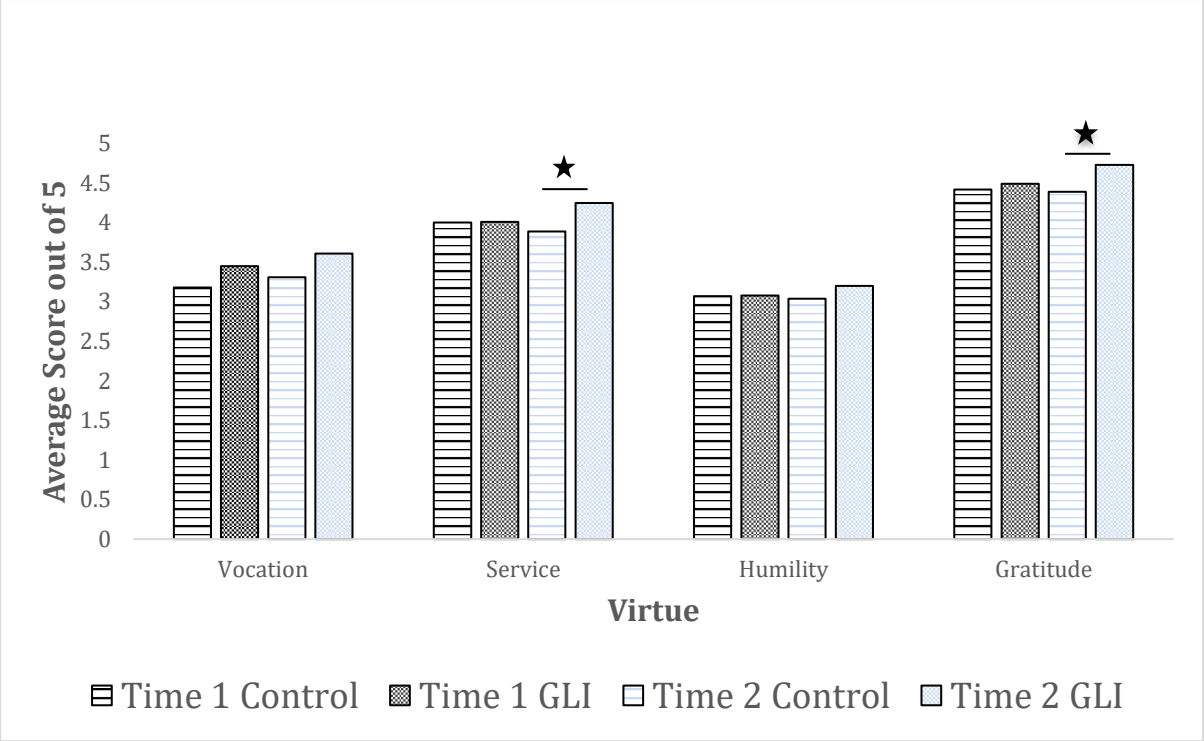
To evaluate differences between the control and GLI groups, we used paired t-tests to compare scores between the two groups at Time 1 and Time 2. For this analysis, in addition to participants that completed surveys at Time 1 and 2, we also included participants who only filled out a survey at Time 1 or 2. This decision was made to increase statistical power given the small sample size. These individuals participated in the full programme but were

not able to complete both pre and post testing. Although we could not evaluate these particular participants' individual change, including their scores enabled us to compare control vs GLI groups as a whole. Thus, we included 9 individuals in the GLI group who solely filled out Time 1 (n = 3) or Time 2 (n = 6) and 8 individuals in the control group who completed the surveys at Time 1 (n = 2) and Time 2 (n = 6). The GLI groups had significantly higher scores at Time 2 compared to the control groups for two virtues: Service and Gratitude (see Table 5 and Figure 1). We did not find any other significant differences for the two other virtues at Time 2, nor for any virtues between groups at Time 1.

Table 5. Comparison of averages scores for the control and GLI groups across Time 1 and 2

	Control *M (SD)*	GLI *M (SD)*	*t* score	*p* value	Effect size
Time 1					
Service	4.00 (.480)	4.01 (.448)	.11	.914	.03 (.01)
Gratitude	4.42 (.666)	4.49 (.491)	.47	.638	.12 (.06)
Humility	3.07 (.341)	3.08 (.321)	.19	.853	.05 (.02)
Vocation	3.18 (.802)	3.45 (.736)	1.31	.197	.34 (.17)
Time 2					
Service	3.89 (.508)	4.25 (.482)	2.44	**.018**	.69 (.33)
Gratitude	4.39 (.612)	4.73 (.336)	2.61	**.012**	.74 (.35)
Humility	3.04 (.293)	3.20 (.334)	1.68	.100	.48 (.23)
Vocation	3.31 (.758)	3.61 (.791)	1.32	.192	.37 (.18)

Figure 1. Average score for each virtue at Time 1 and Time 2 with GLI and control groups



Qualitative Results

Qualitative data drawn from anonymous answers to open-ended questions on the post-test questionnaire illuminate how participants experienced the programme. Short quotations, representative of the wider data set, are used to illustrate how students understood the effects of participation on their cultivation of specific virtues, their moral character, and their career- and life-plans. Each participant is identified through an identification number followed by the year of participation, e.g. [13_2016]. Qualitative data pertinent to the programme's responsiveness to critics of university character education are also offered. The Discussion below makes further use of the qualitative data to illuminate the question of whether the theoretically derived strategy for character development was faithfully implemented. A separate article draws exclusively on the qualitative data to show how the GLI responds to various forms of diversity and trends within emerging adulthood (Blinded for Review, 2019b).

General Impact on Participants

The data generated by the post-test questionnaire demonstrate that the programme was well received by participants. Across the three years of measurement, 73.81% of participants rated the following statement as 'Totally true,' with 97.62% rating it 'Mostly true' or 'Totally true':

> The Global Leadership Initiative offered a group of postgraduates from various academic disciplines and moral, cultural, and religious traditions the opportunity, tools, and support to develop their character and increase their potential to have a positive impact throughout their lives and careers.

This satisfaction with the programme was also seen in the long-answer responses:

> My GLI experience was transformative: it made me into a much more thoughtful and reflective person; it made me much more careful and deliberate about my actions. [3_2017]

> It has been exciting and enriching. I have really enjoyed being part of this community, learning alongside others with similar concerns/interests. [10_2016]

> I am more at peace with myself. I am more capable of managing my emotions and thoughts. I am carrying with me this perspective of ethics when I make decisions, especially decisions concerning career, friendship and family. [7_2016]

Demonstrating understanding of character development as a long-term project, a number of participants explicitly or implicitly note the limitations of measurement undertaken immediately after the programme's completion, recognising the 'proof of the pudding' [11_2016] would only come in later life. As one wrote, 'I've definitely changed - although it's hard to say exactly how' [7_2015]. Many participants recognised that a seven-month programme is, at best, the launch-pad for the ongoing work of cultivating specific virtues and developing character:

> I have lost a lot of my arrogance and I think I do need to go a long way but GLI's reflections will remain with me in this path to humility. [5_2016]

> I am not sure I have changed in any dramatic capacity yet ... The program got things started, and now it's up to us to continue working to have a positive impact and figuring out how to do that. [10_2016]

These reflections affirm both the modesty that Fish encourages about university efforts to shape character (2003b) and the possibility for altering character and life direction that Euben and Kiss emphasize (2010a, 61).

Focal Virtues

Across the three years of measurement, student participants were asked to state to what extent they agreed with the following statement:

> The Global Leadership Initiative has helped me to develop, increase, or strengthen my: Sense of vocation / Commitment to service / Humility / Gratitude

Results for each focal virtue, accompanied by a representative sample of long-answer responses, are given below.

Sense of Vocation:

72.79% of participants agreed or strongly agreed that their sense of vocation had been developed, increased, or strengthened. The long-answer responses support this finding:

> In the beginning of the term I had some doubts about my professional future and how to reconcile it with my vocation. During the two terms, the readings (especially the book on Practical Wisdom), the discussions and self-reflections allowed me to have a clearer understanding of my vocation. [13_2016]

> I am now more than ever reassured about my vocation, and I am aware of the importance of character to follow it successfully through my career. [14_2016]

> I was able to really hone in on what I think my calling is, and to justify that properly. This will be helpful for me in shaping my career and any leadership path I pursue. [3_2017]

Vocation was the only focal virtue where a significant percentage of the students, 11.43%, disagreed or strongly disagreed. This response indicates a continuing degree of uncertainty about vocation, which is reflected in the quotation below and discussed in further detail later in the paper:

> I still feel that the relationship between 'vocation' and my career is not one I can easily understand; it's something I will continue to need to think about. But I think that's part of the point of the programme, which has definitely provided me with some of the tools I'll need to keep doing that--as well as some brilliant insights from my cohort to which I'll continually return. [12_2016]

Commitment to Service:

83.44% of participants agreed or strongly agreed that their commitment to service had been developed, increased, or strengthened:

> I've been reminded that the main driving force behind my desire to lead is my desire to serve and to not lose sight of that. [5_2015]

> I've learned to separate 'service' from 'recognition for service': whereas before I would only consider something to be proper service if many people have heard of it, now I realize that most true service goes mostly unnoticed. This has really affected the way I view my career: I now seek opportunities that don't bring about the most attention but rather those that bring about the most change. [3_2017]

Humility:

82.02% of participants agreed or strongly agreed that their humility had been developed, increased, or strengthened:

> As often emphasised by speakers in the programme, humility is a key aspect of leadership and plays a large role in development of a leader. Humility helps you understand the people you are leading and also makes you open-minded to suggestions and ideas which would help you

> evolve. The practice of humility helps a leader to stay grounded and not be influenced by one's own ego. [6_2015]

> My perspective on humility completely changed after the GLI program. Before, I would think of humility as avoiding taking charge. Now, I think of it as self-knowledge: knowing what one is strong in and acting on that, and realizing when others are stronger in something and allowing them to step forward on that. My perspective developed through the many readings we had that touched on humility, as well as the amazing discussions we had on the topic. [3_2017]

Gratitude:

90.11% of participants agreed or strongly agreed that their gratitude had been developed, increased, or strengthened:

> I've realized how important gratitude is to my well-being and resilience as a leader. [5_2015]

> I have realised how grateful I am! And received some wonderful tools with which to practice gratitude. [8_2016]

Overall, while the participants overwhelmingly perceived that the programme was effective and that they have developed in character and virtue, this was not so clearly shown in the quantitative results. We do see some congruity between the quantitative and qualitative results in the fact that the two focal virtues where we found significant differences between Time 1 and Time 2 and between the GLI groups and the control groups – Service and Gratitude – are also the focal virtues where the participants were most likely to perceive themselves to have developed, increased, or strengthened the virtue. Possible methodological reasons for the discrepancy will be addressed in the Discussion below where the qualitative data is further analysed.

Responsiveness to Critics of University Character Education

Finally, the qualitative data include further evidence for the ways in which the GLI avoids typical criticisms of university character education. First, through the focus on leadership for the common good, the GLI attempts to avoid 'the dangers of distributing moral badges to an already privileged few' (Spelman, 2010, p. 121):

> My professional trajectory has taken a new direction as a result of the programme. I am no longer fixated on pursuing the leadership opportunities that I had once idolized, as ends in themselves. I am, in increasing measure, prepared to serve and inspire others, without the need for title or pomp. I also realize that my uniqueness rests in the combination of successes and failures, which enrich the tapestry of my life. [15_2016]

> Instead of admiring leaders for their achievements or being awed by their rhetorical power, I now attempt to assess them on the basis of more invisible character traits (humility, generosity, gratitude). I used to pride myself on my extensive CV / LinkedIn profile, but now

> I realise that a person is far more than the sum of his or her accomplishments. Actions and gestures speak louder than words. [9_2017]

Second, qualitative data affirm how the GLI eschews 'discipleship' and encourages inclusiveness, dialogue, and self-directed application through the intentional fostering of diversity:

> I valued the fact that the programme fostered exploration rather than being didactic. For me, I have felt that it has been the beginning of an enquiry of deeper reflection on how to better identify my personal morals and apply them in my life and career to have the greatest impact. [2_2015]

> During this time, I gained an appreciation of new conceptions of ethical leadership and crucially, found repeated opportunities to place my own prior[itie]s regarding ethics and leadership, under scrutiny. [15_2016]

> It was amazing discussing various leadership topics with such a diverse group of people - it definitely made me aware of my own ignorance. [4_2016]

> I participated in many programs during my time at Oxford and GLI was no doubt exceptional regarding the level of diverse and intellectually talented students the group brought together for stimulating dialogue. The various groups I participated in, whether academic or athletic, were quite homogeneous in regards to (nationally, race, and possibly other areas) [*sic*]. GLI brought together people from differing backgrounds on a regular basis for deep reflection. [5_2017]

Participants' reflections on dialogue and diversity highlight that the GLI's approach avoids the dangers of uncritical indoctrination and didactic discipleship that worries Fish and other critics.

Discussion

Effects on Individual Participants

In light of these results, there is some evidence in support of the central hypothesis of the GLI. Across the period of measurement, we have seen a general difference between Time 1 and 2 in the intervention group, showing statistically significant improvement in two of the four focal virtues. However, in the control group, there were no statistically significant improvements in any of the four focal virtues. The qualitative data also show participant engagement with all four focal virtues and the perceived impact on their lives and future plans. There are also indications that the approach used insulates the GLI from many of the most prevalent concerns about university character education.

The results elicit two primary questions for discussion: What might explain the fact that intervention groups showed substantial quantitative improvement in only two virtues? And how might future practical activity improve on these results?

First, the improvement in two virtues but not the others might simply be a failure of quantitative measurement owing to inadequacies in the instruments used. For example, a reliability analysis for each virtue found high alphas for both Service (α=.83) and Gratitude (α=.81)— the virtues where improvement was noted—but low alphas for Vocation (α=.49) and Humility (α=.23)—where no significant improvement was shown. This might suggest that not all the items in the instruments intended to measure vocation and humility correlate and reliably test these qualities.

Certainly, it is easy to see how growth in humility might not correspond with rating one's own humility more highly. As Peter Hill and Steven Sandage ask in their analysis of humility measures, 'how do truly humble people report that they are humble?' (2016, p. 137). The instrument used attempts to overcome this challenge, but the very low alpha (.23) indicates it may not have been completely successful.

Similarly, in relation to vocation, the Brief Calling Scale contains two dimensions, one static (e.g. 'I have a calling to a particular kind of work') and the other dynamic (e.g. 'I am searching for my calling as it applies to my career') (Dik et al., 2012, p. 244). The authors of the scale see discerning one's calling as a process, but the two dimensions may actually cut against each other in a longitudinal survey. If one answers affirmatively that they 'have' a calling (the static dimension), they may not answer affirmatively to questions asking whether they are 'searching' for it (the dynamic dimension).

Qualitative data obtained from post-test questionnaires suggest that this may have been true for at least some participants. The person who expressed the least confidence that the GLI programme had contributed to a sense of vocation—'My understanding of vocation did not change much because I came in with a fairly strong religious conception of it. I was very interested to see what my peers who did not share my background had not thought as much about the topic before' [2015_3]—was also the person who, on the quantitative questionnaire, showed the most pronounced increase in their scores. In contrast, another participant who expressed deep appreciation for the GLI's focus on vocation—'I've realized that I should define a vocation rather than a career during my lifetime and while my careers may change over time, my vocation can be my unchanging guide' [2015_5]—showed little change.

Fidelity of Implementation

In addition to measuring effects on individual participants, Berkowitz and Bier note the importance of 'fidelity of implementation' in assessing the effectiveness of character education programmes (2007, p. 41). Did differences in the implementation of pedagogical strategies influence the outcomes, and would greater fidelity to the seven pedagogical strategies that underlie the GLI programming (and habituation through practice in particular) lead to larger effects?

It is clear from the qualitative data that participants are cognizant of two distinct aspects of the programme's approach—the reflective and the practical. Below are indicative quotes with key phrases in bold:

> For me, I have felt that it has been the beginning of **an enquiry of deeper reflection [reflective]** on how to better identify my personal morals and **apply them in my life and career to have the greatest impact [practical].** I especially valued the emphasis throughout GLI on **practices to work to implement the ideas brought up in discussion more rigorously in our lives [practical].** [2_2015]

> The exposure to various key character virtues has made to **think carefully about myself [reflective]** and **develop practices to inculcate them [practical].** [6_2015]

> Every time we met with a guest speaker, I would **become really inspired** to also aim high and do something big**.** I would **note qualities in that person** that I liked **[reflective],** and **I would plan exercises to bring about similar qualities in myself [practical]**. [3_2017]

Yet, while participants seemed to understand the importance of both reflection and practice, further analysis of the qualitative data, involving coding segments of text where participants seemed to be identifying a 'mechanism' for the impact of the programme on their lives, showed interesting differences with respect to Service and Gratitude—the focal virtues where substantial improvement was measured—and Vocation and Humility—the focal virtues that showed no significant change.

Participant responses concerning Service and Gratitude frequently include both the reflective and the practical, as in the following quotes concerned with Gratitude:

> Something very concrete and tangible I picked up from the GLI program is the practice of regular reflection. I now regularly take a notepad and write out my thoughts: my fears, my hopes, my plans, my goals—and also my feelings of gratitude. Listing things for which I'm grateful has made me so much more aware of the privileges I have. [3_2017]

> Started practicing saying thanks to people around me more frequently. [4_2017]

> I have explicitly thanked people more often. [9_2017]

> I thought the session on gratitude was excellent. I really enjoyed the readings, suggested practices, and the discussion. I will seek to incorporate these practices into more of my daily life. [9_2016]

> I knew from previous experience how important the practice of gratitude is in everyday life, but since coming to Oxford I have neglected this practice. Our focus on it in OCP helped me to realise again how important it is, and that I need to re-focus on it again and to develop my own personal practice (like journaling) to sustain gratitude. [7_2016]

In contrast, all the participant comments relating to Vocation and Humility remained in the conceptual, reflective space and were never operationalised into practice. Actions in daily life are never mentioned. For example:

> The programme has made me to ponder over the vocation I wish to pursue in my life. [7_2015]
>
> I have thought about it [vocation] more but still don't have a definition for it. [11_2016]
>
> I thought the vocation discussion we had was the most interesting discussion. It left me with more questions than answers (which I think is good!) but I would have loved to have another discussion of this. [10_2017]

While participants might come to greater appreciation of a virtue and express this positive assessment in their responses, the qualitative data highlights the further step (practice) required to actually develop the virtue. Thus, the results suggest that future work might incorporate particular practices alongside cognitive and reflective engagement to help participants develop the most conceptually complex virtues.

Limitations

This study faces a number of specific limitations. Most theories of character development would lead us to expect any measured character change to be more modest in adulthood than in childhood or adolescence (Grant, 2010, p. 286; Thompson, 2015, pp. 279-306), even when intentional effort is applied (McAdams, 2015, pp. 321–7; Russell, 2015, p. 17; Roberts, Walton, & Viechtbauer, 2006, especially p. 21). This is one component of the criticisms of university character education programmes discussed above (Carr, 2017, pp. 113-4). Although character and virtue development has been considered a whole life's work ever since Aristotle (1999, 1098a16-21), and although there is good reason, supported by studies discussed above, to consider the life-stage of postgraduate students to be a transformative phase of moral formation and value determination, we would still expect the change in adults to be of lesser amplitude than that of children or adolescents. The measurement project hypothesis reflects this by proposing any change should be expected to be 'modest.'

The study is also limited by a small data set. The GLI's theoretically derived strategy is to work intensively with small learning communities of postgraduate students over extended periods of time, and the purpose of this measurement project is to assess the programme's impact. All 41 intervention group participants engaged in the measurement process, but procedural and logistical challenges resulted in 78% of participants offering complete data sets. The sample size precludes the quantity of data required to extrapolate with confidence to larger populations. However, this early-stage study of the first three years of the GLI (October 2014 – September 2017) might, in time, be aligned with other studies, which could increase its contribution to the wider discourse. It would also be useful to re-test cohorts after five years to establish whether the impact persists beyond the end of the programme, particularly since, to date, there is conflicting evidence as to the sustained impact of character development programmes (Berkowitz & Bier, 2007, p. 41–3; Ellenwood, 2014, p. 16).

Given that the measurement relies on self-report, it might be argued that what has been measured is growth in the social desirability of these virtues through participation in the programme rather than growth in the virtue itself. The GLI's theory of moral development

actually recognises the importance of social desirability because empirical studies show that increasing the salience of moral norms is an effective way of cultivating character and encouraging moral behaviour (Miller, 2014, pp. 232–233; Blinded for Review, 2019a). Nonetheless, future research might include a social-desirability scale questionnaire (e.g. Marlowe–Crowne) to control for social desirability, noting whether the results of participants with high scores on the social desirability scale differ from participants with low scores (Exline & Hill, 2012, pp. 210–11).

In addition to improving future programming, this early-stage evaluation also suggests possible avenues for further research. For example, it would be informative to analyse the highly diverse cultural make-up of GLI cohorts, using methods appropriate to cross-cultural research (Goodwin, 1996) to investigate whether there are commonalities or differences in the impact of the intervention within particular religious traditions, ethnicities, or nationalities. In addition, the life-stage of GLI participants might provide an opportunity to study how emerging adults understand the development of vocation in relation to the way they narrate the course of their lives so far. If Colette Daiute (2014) is correct in assuming that 'narrating is a sense-making process—a process for figuring out what's going on in the world and how one fits,' then encouraging participants to narrate their lives would supply precisely the kind of practical activity that might help them to develop their sense of vocation (p.15; cf. Bold, 2012; McAdams, 2015, pp. 321–7).

Conclusion

This article has analysed one effort to cultivate virtue in a university context. It has sought to make two primary contributions to the larger field of character education. First, it has highlighted the value and importance of implementing and assessing character education programmes for postgraduate students within a university context, a population that is largely neglected in a field that tends to prioritize interventions and assessments for children and adolescents. Second, it has sought to show how a voluntary, extracurricular, and intentionally diverse and dialogical character education programme can avoid the most forceful concerns of critics who worry that such efforts are inevitably didactic, indoctrinating, and ineffective within the university context.

In particular, this early-stage evaluation offers quantitative evidence for the modest improvement expected by the GLI with respect to two of the four focal virtues measured (Service and Gratitude), giving grounds for ongoing research and programme development in virtue cultivation amongst postgraduate students. Qualitative data on the four virtues further suggests that participants may have also increased on the other two virtues, though to a lesser extent than Service and Gratitude. Moreover, the analysis of the quantitative data in light of the qualitative data suggests that future iterations of such programming might achieve greater impact if closer attention were paid to ensuring that participants are given not only reflective, conceptual tools but also practical, everyday tasks that may be used in the cultivation of each virtue. Such iterations might further improve our capacity to cultivate and measure virtues in postgraduates.

REFERENCES:

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

Annas, J., Narvaez, D., & Snow, N. E. (Eds.) (2016). *Developing the Virtues: Integrating Perspectives*. Oxford: Oxford University Press.

Anscombe, G. E. M. (1958). Modern Moral Philosophy. *Philosophy,* 33 (124), 1–19.

Aristotle, (1999). *Nicomachean Ethics* (2nd ed.; T. Irwin, Trans.). Indianapolis, IN: Hackett.

Arnett, J. J. (2014a). Presidential Address: The Emergence of Emerging Adulthood: A Personal History. *Emerging Adulthood,* 2(3), 155–162.

Arnett, J. J. (2014b). *Emerging Adulthood: The Winding Road from the Late Teens through the Twenties* (2nd ed.). New York, NY: Oxford University Press.

Author Blinded for Review. (2019a).

Author Blinded for Review. (2019b).

Batson, C. D. (2011). *Altruism in Humans*. New York, NY: Oxford University Press, 2011.

Berkowitz, M. W., & Bier, M. C. (2007). What Works in Character Education. *Journal of Research in Character Education,* 5(1), 29–48.

Berkowitz, M. W. (2011). What Works in Values Education. *International Journal of Educational Research,* 50(3), 153–158.

Bloom, H. S., Michalopoulos, C., & Hill, C. J. (2005). Using Experiments to Assess Nonexperimental Comparison- Group Methods for Measuring Programme Effects. In C. S. Bloom (Ed.), *Learning More from Social Experiments: Evolving Analytic Approaches* (pp. 173-235). New York, NY: Russell Sage.

Bold, C. (2012). *Using Narrative in Research.* London: Sage.

Bowlin, J. R. (1999). *Contingency and Fortune in Aquinas's Ethics*. Cambridge: Cambridge University Press.

Bulach, C. R. (2002). Implementing a Character Education Curriculum and Assessing Its Impact on Student Behavior. *The Clearing House,* 76(2), 79–83.

Burnyeat, M. F. (1980). Aristotle on Learning to Be Good. In A. O. Rorty (Ed.), *Essays on Aristotle's Ethics* (pp. 69–92). Berkeley, CA: University of California Press.

Card, N. A. (2017). Methodological Issues in Measuring the Development of Character. *Journal of Character Education*. 13(2), 29-45.

Carlo, G., Pytlikzillig, L. M., Roesch, S. C., & Dienstbier, R. A. (2009). The Elusive Altruist: The Psychological Study of the Altruistic Personality. In: D. Narvaez & D. K. Lapsley (Eds.) *Personality, Identity, and Character: Explorations in Moral Psychology*. (pp. 271-94) Cambridge: Cambridge University Press.

Carr, D. (2017). Virtue and Character in Higher Education. *British Journal of Educational Studies*, 65(1), 109–124.

Chen, H. T. (2015). *Practical Program Evaluation: Theory-Driven Evaluation and the Integrated Evaluation Perspective* (2nd ed.). London: Sage.

Colby A. (2002) Whose Values Anyway? In: W. Damon (Ed.) *Bringing in a New Era in Character Education* (pp. 149–172). Stanford, CA: Hoover Institution Press.

Crano, W. D. (1996). Quasi-Experimental Designs. In A. S. R. Manstead & M. Hewstone (Eds.), *The Blackwell Encyclopedia of Social Psychology* (p. 464). Oxford: Blackwell.

Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). London: Sage.

Daiute, C. (2014). *Narrative Inquiry: A Dynamic Approach.* London: Sage.

Deutsch, N. L. (2017). Construction and Context: A Response to Methodological Issues in Studying Character. *Journal of Character Education*. 13(2), 53-63.

DeVellis, R. F. (2012). *Scale Development: Theory and Applications* (3rd ed.). London: Sage.

Duffy, R. D., & Sedlacek, W. E. (2007). The presence of and search for a calling: Connections to career development. *Journal of Vocational Behavior*, 70, 590–601.

Dik, B. J., & Duffy, R. D. (2009). Calling and vocation at work: Definitions and prospects for research and practice. *The Counseling Psychologist*, 37, 424–450

Dik, B. J., Eldridge, B. M., Steger, M. F., & Duffy, R. D. (2012). Development and Validation of the Calling and Vocation Questionnaire (CVQ) and Brief Calling Scale (BCS). *Journal of Career Assessment*, 20 (3), 242–63.

Doris, J. (2002). *Lack of Character: Personality and Moral Behavior.* Cambridge: Cambridge University Press.

Elliott, J. C. (2010). *Humility: Development and analysis of a scale* (Doctoral dissertation). University of Tennessee, Knoxville, TN.

Ellenwood, S. (2014, January). *Measuring Virtue Better: A Little Later and a Little Rougher*. Paper presented at the second annual conference of the Jubilee Centre for Character and Virtues, Oxford, UK.

Emmons, R. A., & M. E. McCullough. (2004). *The Psychology of Gratitude*. Oxford: Oxford University Press.

Exline, J. J., & Hill, P. C. (2012). Humility: A consistent and robust predictor of generosity. *The Journal of Positive Psychology*, 7(3), 208–218.

Fish, S. (2003a). Save the World on Your Own Time. *Chronicle of Higher Education*. January 23. Available at: https://www.chronicle.com/article/Save-the-World-on-Your-Own/45335

Fish, S. (2003b). Aim Low. *Chronicle of Higher Education*. May 16. Available at: https://www.chronicle.com/article/Aim-Low/45210

Fish, S. (2004). Why We Built the Ivory Tower. *The New York Times*. May 21. Available at: https://www.nytimes.com/2004/05/21/opinion/why-we-built-the-ivory-tower.html.

Fleeson, W. (2001). Toward a Structure- and Process-Integrated View of Personality: Traits and Density Distributions of States. *Journal of Personality and Social Psychology*, 80(6), 1011–27.

Fox, C., Grimm, R., & Caldeira, R. (2017). *An Introduction to Evaluation*. London: Sage.

Goodwin, R. (1996). A brief guide to cross-cultural psychological research. In J. Haworth (Ed.), *Psychological Research: Innovative Methods and Strategies* (pp. 78-91). London: Routledge.

Grant, R. W. (2010). Is Humanistic Education Humanizing? In E. Kiss & J. P. Euben (Eds.), *Debating Moral Education: Rethinking the Role of the Modern University* (pp. 283-295). Durham, NC: Duke University Press.

Harman, G. (2000). The Nonexistence of Character Traits. *Proceedings of the Aristotelian Society*, 100(1), 223–226.

Harrington, N. G., Giles, S. M., Hoyle, R. H., Feeney, G. J., & Yungbluth, S. C. (2001). Evaluation of the All Stars character education and problem behavior prevention program: effects on mediator and outcome variables for middle school students. *Health Education Behaviour*, 28(5), 533-46.

Hauerwas, S. (1994). *Character and the Christian Life*. Notre Dame, IN: University of Notre Dame Press.

Hill, P. C., & Sandage, S. J. (2016). The promising but challenging case of humility as a positive psychology virtue. *Journal of Moral Education*, 45(2), 132–146.

Hilliard, L. J., Hershberg, R. M., Wang, J., Bowers, E. P., Chase, P. A., Champine, R. B., … Lerner, R. M. (2014). Program Innovations and Character in Cub Scouts: Findings from Year 1 of a Mixed-Methods, Longitudinal Study. *Journal of Youth Development*, 9(4), 6–30.

Hutchins, S. S., Brown, C., Mayberry, R., & Sollecito, W. (2015). Value of a small control group for estimating intervention effectiveness: results from simulations of immunization effectiveness studies. *Journal of Comparative Effect Research*, 4(3), 227–238.

Kiss, E. & Euben, J. P. (2010a). Aim High: A Response to Stanley Fish. In E. Kiss & J. P. Euben (Eds.), *Debating Moral Education: Rethinking the Role of the Modern University*, 57-75. Durham, NC: Duke University Press.

Kiss, E., & Euben, J. P. (2010b). Debating Moral Education: An Introduction. In E. Kiss & J. P. Euben (Eds.), *Debating Moral Education: Rethinking the Role of the Modern University*, 3-26. Durham, NC: Duke University Press.

Kristjánsson, K. (2014). There Is Something About Aristotle: The Pro and Cons of Aristotelianism in Contemporary Moral Education. *Journal of Philosophy of Education*, 48(1), 48–68.

Kristjánsson, K. (2015). *Aristotelian Character Education*. London: Routledge.

Lamb, S., & Randazzo, R. (2016). An examination of the effectiveness of a sexual ethics curriculum. *Journal of Moral Education*, 45(1), 16–30.

Leming, J. S. (2000). Tell Me a Story: An evaluation of a literature based character education programme. *Journal of Moral Education*, 29(4), 413–427.

Lovat, T. (2017). Values education as good practice pedagogy: Evidence from Australian empirical research. *Journal of Moral Education*, 46(1), pp. 88–96.

MacIntyre, A. (1981). *After Virtue: A Study in Moral Theory*. London: Duckworth.

McAdams, D. P. (2015). Psychological Science and the Nicomachean Ethics: Virtuous Actors, Agents, and Authors. In N. E. Snow (Ed.), *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology* (pp. 307–36). Oxford: Oxford University Press.

McCullough, M. E., Emmons, R. A., & Tsang, J. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology*, 82(1), 112–127.

Miles, M., & Huberman, M. *Qualitative Data Analysis: An Expanded Sourcebook* (2nd ed.). London: Sage.

Miller, C. B. (2013). *Moral Character: An Empirical Theory.* Oxford: Oxford University Press.

Miller, C. B. (2014). *Character and Moral Psychology*. Oxford: Oxford University Press.

Morgan, B., Gulliford, L., & Kristjánsson, K. (2017). A new approach to measuring moral virtues: The Multi-Component Gratitude Measure. *Personality and Individual Differences*, 107(1), 179–189.

Narvaez, D. & Lapsley, D. K. (2009). *Personality, Identity, and Character: Explorations in Moral Psychology*. Cambridge: Cambridge University Press.

Noftle, E. E. (2015). Character across Early Emerging Adulthood. In C. B. Miller, R. M. Furr, A. Knobel & W. Fleeson (Eds.), *Character: New Directions from Philosophy, Psychology, and Theology* (pp. 490–517). Oxford: Oxford University Press.

Peterson, C., & Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook and Classification*. Oxford: Oxford University Press.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies. *Psychological Bulletin*, 132(1), 1–25.

Rushton, J. P., Chrisjohn, R. D., & Fekken, G. C. (1981). The altruistic personality and the self-report altruism scale. *Personality and individual differences*, 2(4), 293–302.

Russell, D. (2015). Aristotle on Cultivating Virtue. In N. E. Snow (Ed.), *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology* (pp. 17–48). Oxford: Oxford University Press.

Sabini, J., & Silver, M. (2005). Lack of Character? Situationism Critiqued. *Ethics*, 115(3), 535–562.

Snow, N. E. (2009). *Virtue as Social Intelligence: An Empirically Grounded Theory*. London: Routledge.

Snow, N. E. (2015). *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology*. Oxford: Oxford University Press.

Snow, N. E., Wright, J. C., & Warren, M. T. (in press). *Understanding virtue: Theory and measurement*. New York, NY: Oxford University Press.

Spelman, E. V. (2010). On the Distribution of Moral Badges. In E. Kiss & J. P. Euben (Eds.), *Debating Moral Education: Rethinking the Role of the Modern University* (pp. 211-21). Durham, NC: Duke University Press.

Sreenivasan, G. (2002). Errors about Errors: Virtue Theory and Trait Attribution. *Mind*, 111(441), 47–68.

Sudman, S., & Bradburn, N. (1982). *Asking Questions: A Practical Guide to Questionnaire Design*. London: Jossey-Bass.

Tangney, J. P. (2000). Humility: Theoretical perspectives, empirical findings and directions for future research. *Journal of Social and Clinical Psychology, 19*, 70–82.

Tangney, J. P. (2009). Humility. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (2nd ed., pp. 483–490). New York, NY: Oxford University Press.

Thompson, R. A. (2015). The Development of Virtue: A Perspective from Developmental Psychology. In N. E. Snow (Ed.), *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology* (pp. 279-306). Oxford: Oxford University Press.

Wang, J., Ferris, K. A., Hershberg, R. M., & Lerner, R. M. (2015). Developmental Trajectories of Youth Character: A Five-Wave Longitudinal Study of Cub Scouts and Non-Scout Boys. *Journal of Youth and Adolescence*, 44(12), 2359–2373.

Was, C. A., Woltz, D. J., & Drew, C. (2006). Evaluating character education programs and missing the target: A critique of existing research. *Educational Research Review*, 1(2), 148–156.