What Does Character Education Mean to Character Education Experts?

A Prototype Analysis of Expert Opinions

Robert E. McGrath[1], Hyemin Han[2], Mitch Brown[3], and Peter Meindl[4]

[1]School of Psychology and Counseling, Fairleigh Dickinson University

[2]Educational Psychology Program, University of Alabama

[3]Department of Psychological Science, University of Arkansas

[4]United States Military Academy, West Point

Author Note

Robert E. McGrath https://orcid.org/0000-0002-2589-5088

Hyemin Han https://orcid.org/0000-0001-7181-2565

Mitch Brown https://orcid.org/0000-0001-6615-6081

Peter Meindl https://orcid.org/0000-0002-2686-3723

We have no conflict of interest to report.

Correspondence concerning this article should be addressed to Robert McGrath, School of Psychology and Counseling, Fairleigh Dickinson University, Teaneck NJ 07666. Email: mcgrath@fdu.edu.

Abstract

Having an agreed-upon definition of character education would be useful for both researchers and practitioners in the field. However, even experts in character education disagree on how they would define it. We attempted to achieve greater conceptual clarity on this issue through a prototype analysis in which the features perceived as most central to character education were identified. In Study 1 ($N = 77$), we asked character education experts to enumerate features of character education. Based on these lists, we identified 30 features. In Study 2 ($N = 101$), experts assessed which features were central to character education through a categorization task. In Study 3 ($N = 166$), we assessed the extent of centrality using scalar items. We conclude by offering practical advice for the development of future character education studies and programs rooted in what is deemed central to such programs.

Keywords: Character education; Prototype analysis; Moral education; Reaction time

What Does Character Education Mean to Character Education Experts?

A Prototype Analysis of Expert Opinions

Interest in the study of character education has grown substantially in recent years. A PsycINFO search for the term "character education" yielded 2,335 results, 77% of which came from the past 20 years; searching Google Scholar yielded 60,700 results, with nearly half coming from the same period. The number of programs developed to enhance character traits has also expanded dramatically. The field was largely founded in Kohlberg's (1966) work during the 1960s, but by the mid-2000s Berkowitz and Bier (2007) were able to identify 39 school-based programs for which empirical evidence was available.

The existence of diverse character programs highlights one of the foundational problems in evaluating and discussing character education: character is a broad concept. In the psychology literature alone, there are at least six different classes of competing definitions of character (Fleeson et al., 2014). These range from the narrow (e.g., character as a specific trait such as morality; Hogan, 1973) to the broad (character as personality in general; Shahrokh et al., 2011). In the field of character education in particular, the term is used in connection with moral traits (e.g., kindness, honesty; Kohlberg & Turiel, 1971), civic traits (e.g., tolerance, civility; Baehr, 2017), "performance" traits (e.g., grit, self-discipline; Lickona & Davidson, 2005), intellectual traits (e.g., curiosity, open-mindedness; Baehr, 2013), or some combination of these (McGrath, in press; National Research Council, 2012; Park et al., 2017).

The absence of a consensual definition creates the potential for confusion about what kinds of programs can and should be thought of as character education (Lapsley & Narvaez, 2006). Some programs that identify themselves as character education prioritize only one of these domains largely to the exclusion of others (e.g., Tough, 2011). Others have suggested that

programs should individualize character targets for each student (Linkins et al., 2015). There is similar heterogeneity across programs in the strategies used to enhance student character (Berkowitz & Bier, 2005), further complicating the task of developing a unified definition for character education.

This eclecticism stands in contrast to programs intended to enhance social and emotional learning (SEL), which often share a set of common features. The Collaborative for Academic, Social, and Emotional Learning has identified five core competencies for SEL programs; recommended strategies for addressing these competencies (though the strategies allow for substantial diversity in approach); and protocols for structuring interventions (e.g., Durlak et al., 2011). As a result, it is potentially easier to identify a program as SEL-oriented and to draw generalizations. An effort to develop a more coherent conception of character education could be similarly advantageous.

McGrath (2018) suggested a prototype approach to defining character education. Prototyping is founded in Rosch's (1978) efforts to document how humans naturally use categorical concepts. Whereas membership in scientific and logical categories are typically based on necessary and sufficient conditions, in natural settings people tend to rely on ideal cases, or prototypes. Candidates for inclusion in a category are categorized based on the degree to which they correspond to that prototype (Niedenthal et al., 1985). The elements of the category share a "family resemblance" as opposed to sharing necessary characteristics (Medin et al., 1987). While logical classification tends to set firm boundaries on inclusion in and exclusion from the category, prototype analysis allows for gradations of membership, and even ambiguities near the borders of a category. For example, from a logical perspective, whales and cats are equally mammalian; from the perspective of prototype analysis, the latter is much more a mammal than

the former. This is because cats' features exhibit characteristics more typical of what is conceptualized as mammalian (e.g., fur, four legs).

To demonstrate the potential for the use of a prototype approach in the context of conceptualizing character education, McGrath (2018) suggested that the following seven features could be considered central to character education as opposed to other forms of non-academic learning: the program (1) is school-based, (2) is structured, (3) addresses specific positive psychological attributes, (4) addresses identity, (5) addresses moral growth, (6) addresses holistic growth, and (7) addresses the development of practical wisdom. The key attribute of the prototype approach is that none of these features is considered essential to perceiving a program as representing character education. However, the more of these features that are present, the more likely it is that the program will be considered character education.

These features were offered by McGrath (2018) as examples to concretize the prototype approach. In fact, there is good reason to question these as an adequate prototype. They omit two of the three target domains that have been identified for character education, namely intellectual and performance skills. They also include a concept, practical wisdom, that is primarily associated with specifically Aristotelian forms of character education.

Based on evidence that prototype matching is thought to represent the general strategy by which people put objects into categories, a research model has been developed for uncovering people's implicit prototypes. For example, Kinsella et al. (2015) used samples of adults and college students to describe a prototype for heroes. More directly related to the current investigation, Lapsley and Lasky (2001) developed a prototype for moral character using college students as participants. This research typically requires several sequential studies. The first involves asking a sample of individuals to list features associated with the concept of interest,

with no restrictions on the number of features provided. These features are then distilled into a smaller set based on redundancy, and this set serves as the focus for subsequent studies (see Fehr, 1988; Walker & Pitts, 1998). One of those subsequent studies typically asks participants to evaluate how central or peripheral each feature is to the concept of interest on a Likert-type scale. There is also usually at least one study that involves a behavioral indicator of how central each attribute is to the concept. For example, in some studies, respondents are asked to categorize features as central or peripheral to the concept, and longer reaction times are thought to be indicative of less central features. These additional studies serve to distinguish between central and more peripheral features of the prototype.

Relevant to the topic of character education, prototype development is not always about folk concepts such as heroes. The prototype approach has also been applied in professional contexts, to model the implicit prototypes experts use in their work. The American Psychiatric Association's (2013) manual for the diagnosis of mental disorders bases classification of pathological states on proximity to a prototype developed by an expert work group. For example, the diagnosis of major depressive disorder in this manual is based on two central features and seven more peripheral symptoms. However, these prototypes were based primarily on consensus in the work group after reviewing the literature on a diagnosis rather than the survey methods typically used for folk concepts.

The present series of studies was conducted to generate a formal character education prototype. Because character education represents a technical concept, these studies relied on expert rather than lay public input. The goal was to create a prototype of character education that would be appropriate for professionals in the field. All three studies were approved by the Institutional Review Board of Fairleigh Dickinson University. All data and source code files are

available via the Open Science Framework project space (https://osf.io/xhzbe/?view_only=

09b3dcaad2ca4f11afaa38f98a9b6802).

Study 1

Study 1 followed standard practices for the initiation of work on a prototype.

Specifically, individuals with presumed expertise in the area of character education and related

topics were given the opportunity to generate features they considered relevant to the concept of

character education.

Participants

Participants were (a) researchers who had recently published in *Journal of Moral

Education* and/or *Journal of Character Education* whose work reflected knowledge of character

education, (b) attendees of the annual conference of the Jubilee Centre for Character and Virtues

over its history, (c) recipients of John Templeton Foundation grants related to virtue, (d)

members of the Association for Moral Education and the American Educational Research

Association's Moral Development and Education Special Interest Group, and/or (e) members of

the Society for Personality and Social Psychology or the Moral Science Network listservs. We

distributed a survey link to potential participants in a mass email.

We received input from 77 respondents, of whom 58.0% ($n = 40$) indicated they were

male, 42.0% ($n = 29$) female. The sample was predominantly White ($n = 53$, 80.3%), followed

by Asian ($n = 6$, 9.1%). The large majority reported holding a doctoral degree ($n = 58$, 84.1%),

while the remaining participants all had master's degrees ($n = 11$). Participants described their

discipline, with entries subsequently clustered into broader categories. The largest group

indicated a background in education ($n = 35$, 53.0%), followed by psychology and

developmental studies ($n = 19$, 28.8%), and philosophy/religion ($n = 7$, 10.6%).

Average age was 49.6 (SD = 14.9). Two items described by Talhelm et al (2015) allowed respondents to estimate their political position on social and fiscal issues on a scale from 1 (*Very Liberal*) to 7 (*Very Conservative*). The mean rating was 2.5 (SD = 1.5) on social issues and 3.4 (SD = 1.7) on fiscal issues, suggesting respondents leaned towards liberalism on both questions.

Procedure

After completing the informed consent page and being instructed on what identifying features of a concept entailed, participants received the following instructions: "What are the things you associate with character education? List as many features as you would like."[1] They were then asked as a follow-up question: "What do you consider the best scientific references defining character education?" Responses to this question were not used in the present study. This was followed by a demographic questionnaire and a debriefing form.

Results

Each response was reviewed and partitioned into differentiable features. This process generated a list of 608 such features the respondents associated with character education, a mean of 7.90 features per respondent. Of these 608 features, 105 were deemed irrelevant to the current investigation (e.g., many people listed the names of specific researchers or funding agencies), resulting in a final list of 503 relevant features (a mean of 6.53 features per respondent). Each of the four authors then reviewed the list. By grouping semantically equivalent or closely related

---

[1]One reviewer noted these instructions do not distinguish between character education as it *is* versus character education as it *ought* to be. We do not consider this distinction relevant in the context of prototype analysis, in that the goal is to identify features on which people judge whether a program is an exemplar of character education regardless of the degree to which programs actually demonstrate that feature. Other prototype development studies related to morally tinged constructs have similarly ignored this distinction (e.g., Kinsella et al., 2015; Lapsley & Lasky, 2001). One example that the results reflect *is* more than *ought* is noted in conclusion 6 in the Discussion.

features, each author identified a list of candidates for inclusion in a final feature list. These

results were then collated to review for consistencies across the authors. Through a process of

discussion and consensus, the list was reduced to 30 common features. The final list appears in

Table 1. This Table also includes information on the frequency with which participants

mentioned each feature. To make results across the studies easier to compare, the table also

provides key statistical results from Studies 2 and 3.

We also analyzed the entire text of the responses using the text analysis program

Linguistic Inquiry and Word Count 2015 (LIWC; Pennebaker et al., 2015b). LIWC compares

each word in a passage to a dictionary of terms representing various psychological themes. Most

of these themes are quantified as a percent of words in the text. For example, to generate the

positive emotion score, the frequency of words or phrases in the text that appear in the positive

emotions lexicon (e.g., *love* and *nice*) is computed, and then divided by the number of words in

the text. Similarly, words such as *ally* and *friend* are tallied and used to compute the percent of

words representing the affiliation theme. By default, LIWC computes 41 different word

categories that are considered psychological, e.g., words having to do with emotions and drives,

though some of these are aggregates of other more specific categories. Our analysis focused on

the 34 more molecular scores.

The most common category was words reflecting positive emotions (13.1%), which

reflected the positive outcomes that tend to be associated with character education. The most

common categories of word after positive emotions had to do with performance (work, 10.8%)

followed by words associated with an intellectual skill (insight, 6.6%). This pattern has

interesting implications for the earlier discussion about the centrality of moral, performance, and

intellectual skills to character education. Unfortunately, LIWC does not include categories

reflecting moral themes, so the text analysis did not identify elements of morality in the responses even though Table 1 indicates a strong moral component. Thus, the participants seem to be echoing literature that identifies moral, performance, and intellectual development as three pillars of character education leading to positive development.

LIWC also generates four summary variables: analytical thinking, clout, authenticity, and emotional tone. Based on prior research with very large datasets of text materials, these four variables are converted into percentiles. For example, a clout score of 80 associated with a piece of text indicates the text exceeds 80% of text samples on that variable. The scores for analytic thinking (92nd percentile), clout (73rd), and tone (99th) were higher than typical for text, while that for authenticity (15th) was substantially lower. These findings suggest the text submitted by the respondents was strongly reflective of logical thought (analytic thinking), expertise and confidence (clout), and positivity (tone). The authenticity score is affected by the frequency of self-referential terms (Pennebaker et al., 2015a), so it is not surprising to find this score was lower than typical in writing.

Study 2

The next two studies focused on clarifying the degree to which each of the 30 features identified could be thought of as a central or peripheral element of a prototype. As noted in the introduction, research programs focusing on the development of prototypes typically attempt to distinguish features of the target concept that are relatively central versus peripheral to the concept. This discrimination is often based on subjective ratings and behavioral data. Study 2 used a common methodology for generating both types of data. Specifically, participants are asked to make a binary judgment of each feature, indicating whether or not it is a central feature of the concept. This binary judgment provides subjective data on whether people tend to see the

feature as central. The amount of time it takes the respondent to make the decision is also worthwhile information, as previous research indicates such judgments are easier to make for more central features of a concept and so are associated with lower mean reaction times (e.g., Fehr et al., 1982).

Participants

We recruited participants using the same method described in Study 1. Though 144 respondents began the task, 43 withdrew after reading the instructions, leaving a final sample of 101. Of these, a small majority was male ($n = 51$, 51.5%). Again, the sample was predominantly White ($n = 82$, 83.7%) followed by Asian ($n = 8$, 8.2%), and had doctoral degrees ($n = 82$, 83.7%). This time the most common discipline was psychology and developmental studies ($n = 53$, 56.4%) followed by education ($n = 18$, 19.1%). Participants were asked if they had also participated in Study 1. Most had not ($n = 52$, 53.1%), and most of the remaining participants were unsure ($n = 32$, 32.7%), suggesting that overlap in the samples likely fell somewhere between 14.3% and 47.0%. Average age was 44.2 ($SD = 12.7$). For the political position items described in Study 1, the mean rating on social issues was 2.6 ($SD = 1.5$), and 3.1 for fiscal issues ($SD = 1.5$), again suggesting a liberal orientation in this sample.

Procedure

Consenting participants were presented with the 30 words or phrases one at a time, and their task was to indicate whether each word or phrase was central to the concept of character education, using the "A" key to indicate *Yes, it is* [central to character education], or the "L" key to indicate *No, it is not* [central to character education]. The instructions for this task were to keep one's fingers on the two keys and press the key they deemed appropriate for each feature as it appeared. Participants were instructed to make these judgments as quickly as possible while

remaining accurate, because both speed and their choices were important to the results of the study. They were also warned that once they pressed "A" or "L," they would be automatically advanced to the next term. Participants were given the example of *Mammal* as a concept, and were told that *fur* is central, warranting participants to press "A" for *Yes*, but *tree* would warrant pressing "L" for *No*. On the next page, participants were asked to place fingers on the "A" and "L" keys and keep them there until the task was complete. After 8 seconds, the first feature was presented.

To reduce practice effects, features were presented in random order (Christensen, 2012). In each case, the term appeared in a header font. Below the term the instructions were repeated to indicate whether that term was considered central to character education by pressing "A" for *Yes* or "L" for *No*. The response and the time in seconds from initial presentation to key press was recorded. After evaluating all 30 features, demographic items were collected, followed by debriefing.

Results

Reaction time data were cleaned by removing outliers, defined as reaction times that were greater than three standard deviations above the mean for that feature (Ferguson & Bargh, 2004). This reduced the number of data points for any one feature to 90-98.

The columns in Table 1 headed *p* and Response Latencies reflect the two sources of data: the proportion of respondents who identified the feature as central to character education, and the amount of time it took respondents to make that judgment. The proportion of respondents classifying a feature as central varied between .31 for religious and .95 for morality. Only youth-oriented and religious were endorsed as central by less than half of respondents, suggesting these as the most peripheral of the features listed.

The Response Latency columns provide the means and standard deviations for the response latencies. In general, the calculated means ranged from .76 to 2.10 seconds. Although this range of 1.34 seconds might seem small, it is actually larger than the ranges of mean reaction times reported in similar previous studies. For example, Fehr et al. (1982) reported a range of mean reaction times of .43 seconds, while the range in Fehr and Russell's (1991) study was .47 seconds. As expected, mean response latency was below the median for nine of the ten features most commonly rated as central, indicating respondents were able to judge the centrality of these features relatively quickly. Response latencies were the shortest for virtue, wisdom, empathy development, and morality, with mean latencies of less than one second.

The next column in Table 1 provides point-biserial correlations between the binary judgment of centrality (where responses of *No* were quantified as 0 and *Yes* as 1) and response latency. As hypothesized, these correlations were generally negative, suggesting a respondent who saw the feature as central was generally quicker in responding. The only positive correlations, suggesting faster responses of *No*, were the three features least commonly identified as central (religious, youth-oriented, and behaviorally oriented). One possible explanation for this anomaly is that these last three were so clearly peripheral to many of the participants that they were also able to respond more quickly to them than to features that were less clearly peripheral. The correlation between the 30 values in the *p* column and the 30 means was -.42, indicating that higher proportions were associated with lower mean response times across the 30 features.

Finally, we examined whether prior participation in Study 1 influenced participants' reaction speed. We used frequentist and Bayesian mixed-effects analysis to test whether there was any non-zero effect of prior participation on the mean reaction speed (Han, Park, & Thoma,

2018; Wagenmakers et al., 2018). The results did not suggest an effect of prior exposure to the task. Methodological details and statistical outcomes are provided in the Supplementary Materials.

Study 3

Another common component of studies deriving a prototype involves asking relevant participants to rate the centrality, or prototypicality, of candidate features on a dimensional scale. Where binary evaluations were appropriate for the response latency task in Study 2, since categorical judgments tend to be more difficult for less central features, a dimensional judgment allows conclusions about the degree of centrality of a feature. The final study addressed this issue.

Participants

We recruited participants using the same strategies outlined in Studies 1 and 2. The task was initiated by 185 individuals of whom 19 withdrew for a sample of 166. Of those responding to demographic data, the number of men and women was equal ($n = 78$, 50%). The sample was again predominantly White ($n = 112$, 72.3%) followed by Asian ($n = 17$, 11.0%). There was an unexpectedly high rate of individuals reporting bachelor's degrees ($n = 49$, 30.2%), equaling the number with doctoral degrees. The most common discipline reported was education ($n = 58$, 37.9%), followed closely by psychology and developmental studies ($n = 57$, 34.3%) and a sizeable number of individuals with a background in philosophy, ethics, or theology ($n = 25$, 16.3%). There were 49 participants who indicated they had participated in Study 1 (31.2%) and 82 who did not (52.2%). There were 44 individuals (28.0%) who indicated they had participated in Study 2 and 85 who had not (54.1%). Average age was 51.3 ($SD = 15.2$). Consistent with

Studies 1 and 2, the mean for the social issues item was 2.8 ($SD = 1.4$), for the fiscal item 3.5

($SD = 1.4$).

Procedure

Consenting respondents were presented with each of the 30 features with the following

instructions: "Please use the scales below to indicate how much each term is central to character

education." They rated each on a scale from 1 (*Not At All Central* [to character education]) to 7

(*Extremely Central*). The order of presentation was randomized across participants. After rating

all 30 features, participants completed demographic questions and were debriefed.

Results

The means and standard deviations for the ratings may be found in the Centrality Ratings

columns of Table 1. The means fall in almost exactly the same order as the proportions in Study

2, with morality rated as the most central and religious as least central feature. In fact, the

correlation across the 30 features between the proportion identifying the feature as central in

Study 2 and the mean centrality rating in Study 3 is .94. The only noticeable discrepancy occurs

for nurturing and socio-cultural support, for which the means fell somewhat lower than expected

from the proportions in Study 2. Given the high rate of consistency in rank ordering, it is not

surprising to find that the correlation between mean reaction time in Study 2 and mean centrality

rating in Study 3 is -.38, almost the same as that found in Study 2 between proportion and mean

reaction time (-.42).

We wanted to examine the consistency between participants' responses between Studies

2 and 3, but given that respondents did not identify themselves, we were limited to those

individuals whose data from the two studies we could match based on internet protocol

addresses. This restricted our comparison to 25 participants. We conducted frequentist and

Bayesian mixed-effects logistic regression to check for consistency across the 30 features (Han, Park, & Thoma, 2018; Wagenmakers et al., 2018). As expected, we found evidence of a large effect size suggesting presence of significant consistency (Ferguson, 2009). Methodological details are provided in the Supplementary Materials.

We performed a principal components analysis (PCA) to explore the dimensionality of the features while minimizing information loss (Jolliffe & Cadima, 2016). A parallel analysis using the defaults in the psych package (Revelle, 2019) in R suggested retaining two factors. Promax rotation was implemented, and the correlation between components was estimated at r = .49. Table 2 presents the factor loadings (factor loadings $\geq$ .40 are bolded for reference). Keeping in mind these reflect covariation in participants' ratings of centrality, there seemed to be a tendency to rate features that reflect interpersonal functioning similarly, while ratings of broader themes such as wisdom were similarly more related. There were also five features (i.e., personal, youth-oriented, flourishing, behaviorally oriented, and educational) that did not load highly on either component.

We also conducted a network analysis to explore the relationships between pairs of features (Costantini et al., 2015). Connections between features were modeled using correlations. We lasso regularized the estimated coefficients of the relationships, a procedure for model estimation that penalizes more complex models (Epskamp et al., 2018). To further simplify interpretation, we excluded relationships from the plot when $r < .10$ given that this value is often considered the threshold for a small but non-trivial effect size.

Network analysis allows us to examine the centrality of each feature to the set of features as a whole using the expected influence (EI). A higher EI is indicative of a more influential or central feature in a network (Robinaugh et al., 2016). These statistics were calculated using the

centralityTable function implemented in the qgraph R package (Epskamp et al., 2012). Simultaneously, we also visually inspected the network plot to search for central features. The resulting network plot is presented in Figure 1, while the EIs are reported in Table 3. Emotional growth, caring, and virtue showed the strongest influence on other variables.

Discussion

Centrality Statistics

With the help of relevant experts, we developed a prototype for character education consisting of 30 features. For purposes of interpretation, some researchers divide the list of prototype features they develop into two groups of approximately equal size to indicate the more central and more peripheral elements of the prototype. Using this heuristic, results in Table 1 suggest the entries from morality to relationship development represent more central features of the prototype, with the highest proportion of respondents in Study 2 identifying those as central.

Other results in the table are consistent with this dichotomization. In Study 2, the mean response latency across the 30 features was 1.28 seconds. The mean for the top 15 was 1.15 seconds, while that for the lower 15 was 1.40. Similarly, the mean centrality rating across the 30 features on the 1-7 Likert scale was 5.24. The means for the top 15 all met or exceeded this level, while all of the features listed below relationship development were associated with mean ratings < 5.24.

The entries at the top of the table should probably be unsurprising to most advocates of character education, but several points can be made about specific features. It is interesting to note the feature character strengths was the second entry in the list while character traits were among the more peripheral features. This outcome suggests that despite its relative recency, the VIA Classification of Character Strengths and Virtues (Peterson & Seligman, 2004) has had a

significant effect on how character education is perceived. However, it could also be the case that some respondents perceived the term trait as implying fixed rather than malleable attributes, or more relevant to personality theory than character education. Similarly, though SEL and character education developed as distinct approaches to skills education (Elias et al., 2007), the influence of the former on the latter is evident from the presence of social-emotional learning in the list of more central features. This may also to some extent reflect recent efforts to synthesize the two approaches (Elias, 2014). That said, the relatively low placement of skills development and behaviorally oriented in the list may suggest an implicit commitment to whole-person development in character education versus specific skills development in SEL.

Other terms in the more peripheral list also merit comment. The relatively low standing of educational and youth-oriented suggest an interest in character education across settings and developmental stages. The location of citizenship in the lower group is also worth noting given that some commentators have considered enhanced citizenship a key goal for character education programs (Althof & Berkowitz, 2006). In saying this, it must be remembered the features in the bottom half of Table 1 are not irrelevant to character education; only two are associated with means near or below the neutral point on the centrality rating scale: youth-oriented and religious. The bottome 15 features are only more peripheral to the concept relative to those higher in the list. In contrast, two of the features suggested by McGrath (2018) for a character education prototype—school-based (though educational made the list) and structured—are not reflected in the list, at least not directly.

Finally, it is noteworthy that modeling is the only entry in the list that reflects an implementation strategy, and this falls in the more peripheral set. The dominance of outcomes in the list suggests character education is understood primarily in terms of its goals. A few of the

terms instead describe some general principles of character education (e.g., educational, youth-oriented, personal), but these tend to fall rather low on the list. This emphasis on outcomes could be a result of the instructions of Study 1 not specifying whether respondents describe the supposed outcomes of character educations or the implementation. Though such non-specific instructions for prototype research are commonplace (e.g., Rosch, 1978), these instructions could have precluded us from identifying other aspects of character education beyond the outcomes more readily. Future research would benefit from specifically tasking participants with identifying central features for character education beyond outcomes.

Relationships between Centrality Ratings

The results from the PCA suggested two clusters of centrality ratings. The first set is more indicative of a set of goals reflecting effectiveness as a person of character. The second set tends to encompass broader and more abstract concepts addressed in character education programs such as morality and virtue. The relative order of the two components reflects the number of terms in the set that are representative of the two themes, with more terms reflective of personal effectiveness.

The network plot more explicitly visualizes how features are related to each other. Consistent with the results from PCA, two visible clusters can be identified from the plot, one centered around emotional growth/caring and one around virtue. The EIs similarly indicate that emotional growth and caring were the most influential drivers of ratings in this first group, while virtue was particularly influential on ratings in the second group. These findings again suggest that participants were considering the features of character education primarily in two domains, one about interpersonal strengths, the other about more broader features embracing morality and virtues.

These results further support the synthesis of elements from SEL and character education, even while findings discussed above suggest differences in emphasis. Character education seems to share the development of skills with SEL (e.g., emotional management, goal setting and pursuant, empathy, establishing positive relationships, responsible decision making; CASEL, 2020) while also emphasizing the role of broader social themes such as virtues and citizenship (Carr, 2008). Interestingly, flourishing seems to connect these two categories in the network plot. This finding suggests flourishing as a common influencer in both clusters of ratings, a conclusion supported by the PCA finding that flourishing was about equally related to the two components. The concept of flourishing has historically been explicitly connected with virtue concepts (e.g., Kristjánsson, 2012), while Schonert-Reichl (2019) concluded that the cultivation of social and emotional skills can be motivated by the goal of helping individuals in communities flourish.

Limitations and Conclusions

Several limitations of the study are worth noting. There was no attempt to assure a minimum familiarity with character education. In fact, members of the Society for Personality and Social Psychology and the Moral Science Network listservs could have had little familiarity with the field. Because these were professional lists, we assume that individuals relatively unfamiliar with character education would not have participated, though we cannot be sure this was true in all cases. The recruitment strategy may also have biased the response pool towards researchers rather than pure practitioners.

It is also unclear to what extent conversation occurs within the character education community about the meaning of the term. As noted previously, CASEL has provided relatively clear guidance about the desired characteristics of SEL programs. In contrast, the character education field has demonstrated substantial diversity. In the absence of systematic discussion, it

is difficult to gauge how well-formed are people's perceptions of the prototypical character education program. That said, variability in centrality ratings across features does indicate some consensus in the field as a whole.

Finally, we noted there was overlap in the samples, so certain individuals drove the results more than others. That would probably have had its largest impact on reaction times, and as discussed, we found no evidence of a systematic difference across those with or without prior familiarity with the project. The small population of potential respondents also limited our capacity to recruit completely unique samples. Future replications nonetheless could benefit from ensuring the use of unique samples in each study.

We would also note that several responses on Study 1 reflected negative perceptions about character education (e.g., counterproductive). These were so unusual that it was not reasonable to include them in the list of features. It is possible, though, that some individuals experience character education negatively, and this study was insufficient to tap effectively into that mindset. It would be an interesting topic for future research to gauge the extent of negative perceptions of character education, particularly if those criticisms help to develop more effective or cross-culturally valid character education models. This concern is particularly salient when considering the primary demographic profile for our respondents was older White men. Future work would benefit from purposely recruiting character education scholars across varying cultures, ethnicities, and ages, while also ensuring greater representation of women in the sample.

With these limitations in mind, we can draw some conclusions about what experts in character education mean when they use the term, and how character educators can use this information:

1) Given the absence of systematic attempts to develop a unifying definition of character education, we believe the prototype approach is particularly useful as a means of conceptualizing what characterizes a character education program. The prototype strategy eschews clear dividing lines between inclusion in and exclusion from the category; instead, it provides a framework for saying one program is more clearly a character education program than another.

2) In that regard, the more features appearing near the top of the list in Table 1, the more prototypical that program is of character education.

3) Educators interested in developing character education programs would similarly be well-served by considering those features at the top of Table 1 for inclusion as targets.

4) Character education seems to be defined by experts in terms of its outcomes, not its methods, treatment providers, or any other element of program process. The only exception to this conclusion is that modeling seems a relevant, though not particularly central, feature of what a character education program.

5) Though many character education programs focus on school-based outcomes such as school behavior (see Berkowitz & Bier, 2007), the defining features of character education have more to do with broad domains of functioning. This of course makes sense in the context of programs intended to modify something as ingrained in the individual as character, and suggests that program outcomes should always be extended beyond the potentially low-hanging fruit of school behaviors to include some of these broader constructs.

6) Of the four broad elements of character development noted earlier—moral, performance, intellectual, and civic—the first seems to be prioritized, as exemplified by the location of terms such as morality and empathy development near the top of the list. Several features

such as self-regulation and deliberation would suggest the performance domain is stressed next. Though some of the features are reflective of insight as noted earlier, such as identity development, features reflecting the intellectual capacity to gather and process information effectively are absent from the list. Despite some strong advocates for the importance of developing intellectual character (e.g., Baehr, 2013) or for simultaneously emphasizing moral, performance, and intellectual development (McGrath, in press), experts in character education seem to place less importance on intellectual development as a target. The civic component is largely absent from the list, suggesting that its association with character education is much more tenuous than the other three. That said, these conclusions may not be as clear as this discussion suggests, since the list includes some features that are broad enough to encompass even all four of these domains, particularly character strengths and virtues.

References

Althof, W., & Berkowitz, M. W. (2006). Moral education and character education: Their relationship and roles in citizenship education. Journal of Moral Education, 35(4), 495-518. https://doi.org/10.1080/03057240601012204

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: Author.

Baehr, J. (2013). Educating for intellectual virtues: From theory to practice. *Journal of Philosophy of Education, 47,* 248-262. https://doi.org/10.1111/1467-9752.12023

Baehr, J. (2017). The varieties of character and some implications for character education. *Journal of Youth and Adolescence*, *46(6)*, 1153-1161.

Berkowitz, M. W., & Bier, M. C. (2005). *What works in character education: A research-driven guide for educators*. Washington, DC: Character Education Partnership.

Carr, D. (2008). Character education as the cultivation of virtue. In L. P. Nucci & D. Narvaez (Eds.), Handbook of moral and character education (pp. 99-116). Routledge.

CASEL. (2020). *What is SEL*. Available at http://www.casel.org/what-is-sel

Christensen, L. (2012). Types of designs using random assignment. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology: Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (p. 469–488). American Psychological Association.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality, 54,* 13-29. https://doi.org/10.1016/j.jrp.2014.07.003

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development, 82,* 405-432. https://doi.org/10.1111/j. 1467-8624.2010.01564.x

Elias, M. J. (2014). The future of character education and social-emotional learning: The need for whole school and community-linked approaches. *Journal of Character Education, 10*(1), 37-43.

Elias, M. J., Parker, S. J., Kash, V. M., & Dunkeblau, E. (2007). Social-emotional learning and character and moral education in children: Synergy or fundamental divergence in our schools? *Journal of Research in Character Education, 5*(2), 167-181.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*(1), 195-212. https://doi.org/ 10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software, 48*(4), 1-18. https://doi.org/10.18637/jss.v048.i04

Fehr, B. (1988). Prototype analysis of the concepts of love and commitment. *Journal of Personality and Social Psychology, 55*(4), 557–579. https://doi.org/10.1037/0022-3514.55.4.557

Fehr, B., & Russell, J. A. (1991). The concept of love viewed from a prototype perspective. *Journal of Personality and Social Psychology, 60*(3), 425-438. https://psycnet.apa.org/doi/10.1037/0022-3514.60.3.425

Fehr, B., Russell, J. A., & Ward, L. M. (1982). Prototypicality of emotions: A reaction time

      study. *Bulletin of the Psychonomic Society, 20*(5), 253-254. https://doi.org/10.3758/

      BF03334831

Ferguson, C. J. (2016). An effect size primer: A guide for clinicians and researchers. In A. E.

      Kazdin (Ed.), *Methodological issues and strategies in clinical research* (pp. 301–310).

      American Psychological Association. https://doi.org/10.1037/14805-020

Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: The effects of goal pursuit on

      automatic evaluation. *Journal of Personality and Social Psychology, 87*(5), 557-572.

      http://doi.org/10.1037/0022-3514.87.5.557

Fleeson, W., Furr, R. M., Jayawickreme, E., Meindl, P., & Helzer, E. G. (2014). Character: The

      prospects for a personality-based perspective on morality. *Social and Personality*

      Psychology Compass, *8*(4), 178-191.

Han, H., Park, J., & Thoma, S. J. (2018). Why do we need to employ Bayesian statistics and how

      can we employ it in studies of moral education? With practical guidelines to use JASP for

      educators and researchers. *Journal of Moral Education, 47*(4), 519-537.

      https://doi.org/10.1080/03057240.2018.1463204.

Hogan, R. (1973). Moral conduct and moral character: a psychological perspective.

      Psychological Bulletin, 79(4), 217-232.Jolliffe, I. T., & Cadima, J. (2016). Principal

      component analysis: A review and recent developments. *Philosophical Transactions of*

      the Royal Society A: Mathematical, Physical and Engineering Sciences, 374(2065),

      20150202. https://doi.org/10.1098/rsta.2015.0202

Kinsella, E. L., Ritchie, T. D., & Igou, E. R. (2015). Zeroing in on heroes: A prototype analysis of hero features. *Journal of Personality and Social Psychology, 108*(1), 114-127. https://doi.org/10.1037/a0038463

Kohlberg, L. (1966). Moral education in the school: A developmental view. *School Review, 74*(1), 1-30.

Kristjánsson, K. (2012). Positive psychology and positive education: Old wine in new bottles? *Educational Psychologist, 47*(2), 86-105. https://doi.org/10.1080/00461520.2011.610678

Lapsley, D. K., & Lasky, B. (2001). Protoypic moral character. *Identity: An International Journal of Theory and Research, 1*(4), 345-363. https://doi.org/10.1207/S1532706XID0104_03

Lapsley, D. K., & Narvaez, D. (2006). Character education. In A. Renninger & I. Siegel (Eds.) and W. Damon & R. Lerner (Eds.-in-Chief.), *Handbook of child psychology: Vol. 4. Child psychology in practice* (6th ed., pp. 248-296). Wiley.

Lickona, T., & Davidson, M. (2005). *Smart & good high schools: Integrating excellence and ethics for success in school, work, and beyond*. Cortland, NY: Center for the 4th and 5th Rs/Character Education Partnership. Available at https://www2.cortland.edu/centers/character/high-schools/SnGReport.pdf

Linkins, M., Niemiec, R. M., Gillham, J., & Mayerson, D. (2015). Through the lens of strength: A framework for educating the heart. *Journal of Positive Psychology, 10*(1), 64-68. https://doi.org/10.1080/17439760.2014.888581

McGrath, R. E. (2018). What is character education? Development of a prototype. *Journal of Character Education, 14*(2), 23-35.

McGrath, R. E. (in press). Darwin meets Aristotle: Fundamental virtues as contributors to

addressing current cultural tensions. *Journal of Positive Psychology*. http://doi.org/10.

1080/17439760.2020.1752781

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual

cohesiveness, and category construction. *Cognitive Psychology, 19*(2), 242-279.

http://doi.org/10.1016/0010-0285(87)90012-0

National Research Council. (2012). *Education for life and work: Developing transferable

knowledge and skills in the 21st century*. National Academies Press.

Niedenthal, P. M., Cantor, N., & Kihlstrom, J. F. (1985). Prototype matching: A strategy for

social decision making. *Journal of Personality and Social Psychology, 48*(3), 575–584.

https://doi.org/10.1037/0022-3514.48.3.575

Park, D., Tsukayama, E., Goodwin, G. P., Patrick, S., & Duckworth, A. L. (2017). A tripartite

taxonomy of character: Evidence for intrapersonal, interpersonal, and intellectual

competencies in children. *Contemporary Educational Psychology, 48,* 16-27. https://doi.

org/10.1016/j.cedpsych.2016.08.001.

Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015a). *Linguistic Inquiry and

Word Count: LIWC2015*. Pennebaker Conglomerates.

Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015b). *The development and

psychometric properties of LIWC2015*. University of Texas at Austin.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A classification and

handbook*. American Psychological Association/Oxford University Press.

Revelle, W. (2019). psych: Procedures for psychological, psychometric, and personality research (Version 1.9.7) [Computer software]. Retrieved from https://cran.r-project.org/web/packages/psych/

Robinaugh, D. J., Millner, A. J., & McNally, R. J. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of Abnormal Psychology, 125*(6), 747-757. https://doi.org/10.1037/abn0000181

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorisation* (pp. 27-48). Erlbaum.

Shahrokh, N. C., Hales, R. E., Phillips, K. A., & Yudofsky, S. C. (2011). The language of mental health: A glossary of psychiatric terms. *American Psychiatric Association*.

Schonert-Reichl, K. A. (2019). Advancements in the landscape of social and emotional learning and emerging topics on the horizon. *Educational Psychologist, 54*(3), 222-232. https://doi.org/10.1080/00461520.2019.1633925

Talhelm, T., Haidt, J., Oishi, S., Zhang, X., Miao, F. F., & Chen, S. (2015). Liberals think more analytically (more 'WEIRD') than conservatives. *Personality and Social Psychology Bulletin, 41*(2), 250-267. doi:10.1177/0146167214563672

Tough, P. (2011, September 18). What if the secret to success is failure? *New York Times Magazine*. Available at http://www.nytimes.com/2011/09/18/magazine/what-if-the-secret-to-success-is-failure.html

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., ... & Meerhoff, F. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*(1), 58-76. https://doi.org/10.3758/s13423-017-1323-7

Walker, L. J., & Pitts, R. C. (1998). Naturalistic conceptions of moral maturity. *Developmental Psychology, 34*(3), 403-419. https://doi.org/10.1037/0012-1649.34.3.403

Table 1

Character Education Features.

Feature	f (%)	p	Response Latencies M	SD	r	Centrality Ratings M	SD
Morality	301 (13.66%)	.95	0.99*	0.81	-.36	6.17	1.17
Character strengths	255 (11.57%)	.92	1.05*	0.72	-.39	5.85	1.37
Empathy development	38 (1.72%)	.92	0.91*	0.44	-.26	5.81	1.37
Virtue	415 (18.83%)	.91	0.76*	0.54	-.32	5.87	1.33
Self-regulation	44 (2.00%)	.89	1.18*	0.80	-.15	5.82	1.26
Caring	4 (.18%)	.89	1.24	1.33	-.33	5.59	1.40
Emotional growth	8 (.36%)	.87	1.12*	0.69	-.19	5.58	1.35
Positive development	151 (6.85%)	.86	1.18*	0.87	-.19	5.49	1.37
Wisdom	26 (1.18%)	.84	0.90*	0.76	-.51	5.47	1.33
Identity development	19 (.86%)	.83	1.32	0.95	-.05	5.42	1.36
Justice	5 (.23%)	.82	1.04*	0.89	-.43	5.66	1.27
Socially focused	35 (1.59%)	.81	1.44	0.97	-.30	5.24	1.49
Character traits	18 (.82%)	.81	1.46	1.06	-.28	5.47	1.59
Social-emotional learning	9 (.41%)	.80	1.43	1.01	-.22	5.54	1.40
Relationship development	83 (3.77%)	.78	1.28	0.87	-.25	5.48	1.36
Positive motivations	12 (.54%)	.78	1.64	1.09	-.37	5.23	1.46
Flourishing	134 (6.08%)	.76	1.12*	0.83	-.32	5.13	1.59
Nurturing	4 (.18%)	.73	1.22*	0.93	-.17	4.93	1.47
Socio-cultural support	77 (3.49%)	.72	1.83	1.32	-.13	4.85	1.51
Deliberation	123 (5.58%)	.70	1.29	0.82	-.01	5.12	1.48
Educational	108 (4.90%)	.69	1.35	0.92	-.31	5.10	1.53
Autonomy	11 (.50%)	.68	1.40	1.01	-.22	5.01	1.50
Modeling	58 (2.63%)	.68	1.17*	0.69	-.16	5.14	1.48
Holistic	4 (.18%)	.67	1.21*	0.72	-.26	5.07	1.60
Citizenship	63 (2.86%)	.65	1.04*	0.69	-.29	5.18	1.56
Personal	13 (.59%)	.61	1.68	1.29	-.12	5.02	1.42
Skills development	20 (.91%)	.57	1.31	0.88	-.20	4.91	1.50
Behaviorally oriented	111 (5.04%)	.56	2.10	2.01	.05	4.82	1.53
Youth-oriented	8 (.36%)	.33	1.53	1.01	.11	4.06	1.73
Religious	47 (2.13%)	.31	1.09*	0.69	.05	3.25	1.86
M		.74	1.28		-.22	5.24	

*M < median (1.23)

Note. f and *%* are the number and percent of descriptors in Study 1 that were grouped into each feature. Some frequencies are very high because individual participants provided multiple variants of the same concept (e.g., *morality*, *ethics*, and *moral principles*). *p* is the proportion of respondents in Study 2 who identified the feature as central. Response latency values are the mean and standard deviation for the number of seconds it took the respondent to make that judgment. *r* is the point-biserial correlation between these two variables. Features are sorted according to the proportion of respondents in Study 2 indicating the feature is central to character education.

Table 2

Factor Loadings of Features in Study 3.

Feature	Interpersonal themes	Broad themes
Empathy development	**0.86**	-0.21
Socially focused	**0.80**	-0.28
Social-emotional learning	**0.80**	-0.17
Emotional growth	**0.76**	-0.02
Caring	**0.76**	-0.03
Positive development	**0.73**	-0.02
Relationship development	**0.72**	-0.11
Socio-cultural support	**0.72**	-0.01
Nurturing	**0.64**	0.06
Holistic	**0.63**	0.00
Positive motivations	**0.62**	0.05
Identity development	**0.58**	-0.18
Justice	**0.56**	0.06
Skills development	**0.55**	-0.03
Citizenship	**0.51**	0.09
Autonomy	**0.46**	-0.03
Self-regulation	**0.41**	0.08
Deliberation	**0.40**	0.11
Virtue	-0.22	**0.92**
Character traits	-0.26	**0.88**
Character strengths	-0.10	**0.73**
Wisdom	-0.04	**0.62**
Religious	-0.22	**0.62**
Morality	0.15	**0.45**
Modeling	0.23	**0.41**
Personal	0.35	0.21
Youth-oriented	0.34	0.12
Flourishing	0.33	0.34
Behaviorally oriented	0.33	0.33
Educational	0.23	0.24
Variance proportion	0.28	0.13

Note. Bolded factor loadings represent loadings ≥ .40.

Table 3

Expected Influence Values for Features in Study 3.

Feature	Expected Influence
Emotional growth	2.17
Caring	1.70
Virtue	1.35
Positive development	1.31
Empathy development	1.26
Holistic	1.10
Socio-cultural support	0.86
Social-emotional learning	0.66
Socially focused	0.58
Nurturing	0.44
Justice	0.40
Behaviorally oriented	0.40
Relationship development	0.19
Positive motivations	0.15
Character traits	-0.06
Flourishing	-0.20
Skills development	-0.20
Citizenship	-0.31
Personal	-0.56
Character strengths	-0.61
Wisdom	-0.67
Modeling	-0.72
Self-regulation	-0.84
Deliberation	-0.91
Autonomy	-1.04
Youth-oriented	-1.07
Identity development	-1.13
Morality	-1.21
Educational	-1.36
Religious	-1.67

Figure 1

Network Analysis of Centrality Ratings in Study 3. BO=Behaviorally oriented; CS=Character strengths; CT=Character traits; Delib=Deliberation; Education=Educational; EG=Emotional growth; ED=Empathy development; Flourish=Flourishing; ID=Identity development; PD=Positive development; PM=Positive motivations; RD=Relationship development; Rel=Religious; SR=Self-regulation; SD=Skills development; SEL=Social-emotional learning; SF=Socially focused; SCS=Socio-cultural support; YO=Youth-oriented.