

Teacher Perceptions of Assessments in Character Education: A Case Study

Sarah Hickman

A Senior Thesis submitted in partial fulfillment
of the requirements for graduation
in the Honors Program
Liberty University
Fall 2019

Acceptance of Senior Honors Thesis

This Senior Honors Thesis is accepted in partial fulfillment of the requirements for graduation from the Honors Program of Liberty University.

Monica Huband, Ed.D.
Thesis Chair

Scott Watson, Ph.D.
Committee Member

James H. Nutter, D.A.
Honors Director

Date

Abstract

Since the time of Aristotle, educators have emphasized character as a necessary part of a student's education, and currently, many states mandate character education by law. Because of this historical and legal emphasis, there is a growing discussion on the necessity of assessments in character education to ensure that character education programs are effective. While there is research on the large-scale effectiveness of programs with different assessments, there is little research about how teachers perceive assessments and measurements in character education. This study was conducted to begin to address this gap in the research. Through the research design of an embedded single case study, the researcher sought to answer the questions of how teachers perceive assessments in character education and how their answers differed when speaking about different grade levels. Data were collected from semi-structured interviews with six different teachers and were analyzed through methods of grounded theory. The findings suggest that teachers believe that measurements for character would be a helpful tool, but they are concerned about the practicality and plausibility of assessments. These findings should encourage researchers, policymakers, and administrators to consider the voices and concerns of teachers while reviewing other research about authentic and effective methods for character education assessment.

Teacher Perceptions of Assessment in Character Education: A Case Study

Since the time of Aristotle, educators have emphasized character as a necessary part of a student's education. This emphasis continues in modern public schools; eighteen states mandate character education by law while another eighteen encourage it through legislation ("Character Education," n.d.). However, there is still the question of accountability. While character education is mandated, the question remains of whether or not the character education that is being taught is effective. If it is effective, then it should be continued. However, if it is not effective, schools and states may need to reevaluate character education programs. In subjects like math and English, assessments are given to demonstrate instructional effectiveness and students' growth. Because the law expects teachers to be teaching character to their students, some researchers are beginning to argue for more assessments in character education. For example, in 2017, *The Journal of Character Education* released an entire issue to bring attention to important of using evidence to determine the effectiveness of character education programs (Beatty, 2017). While there is a growing body of research that advocates for assessments and also employs different assessments, there is no qualitative research on how teachers perceive character education assessments. Also, there is no research that examines how views of character assessment differ among various teachers in the same school setting based on grade level. It is important for research to take into account qualitative research in assessments in character education along with quantitative research.

Research Questions

1. How do teachers perceive assessments in character education?
2. How does the grade level that teachers teach impact their perceptions of assessments in character education?

Character Education

Teacher Perceptions

How teachers perceive character education itself is a beginning step to understanding the effectiveness of character education and teacher's perceptions of assessments in character education. While administrative and school-wide changes in character education and assessment are helpful, teachers are the ones who must implement these changes. As Roberston-Kraft and Austin (2015) point out, the teachers are the true mechanisms of change in schools. Though the effectiveness of character education or assessments is often determined by quantitative tests, the qualitative evidence of teacher perceptions are not considered.

In one study done by Leming & Yendol-Hoppey (2004), the perceptions of teachers about character education boiled down to five major themes. These themes were that teachers did not see character education as controversial, think of character education programs as opportunities to become closer to their students, make character education part of their entire curriculum, want to focus on moral decision-making, and see that building character takes a whole community of support (Leming & Yendol-Hoppey, 2004). From this study, it can be inferred that teachers believe that character education is a positive opportunity in school. While they desire to focus on moral decision-making in character education, they did not mention how they might measure students' learning in moral decision-making.

In a qualitative assessment of sixty different interviews with teachers from grades K-12, Meidl (2008) identified similar themes in teachers' perceptions of character. The first was that teachers had a difficult time defining character. The way that many teachers resorted to explaining character was to focus on positive behaviors. They also spoke about how mandating character education was not an effective strategy in light of all the time needed for mandated

standardized tests (Meidl, 2008). Also, some said that they felt like character education was one way the government could try to fix complex social issues with a program (Meidl, 2008).

Teachers want character education to be simple to fit it into their schedules, but also worry about the effects of character education programs that are too simple (Meidl, 2008). Most of character education in elementary classrooms comes through teachable moments where teachers catch students with either positive or negative behaviors and then either reward or discipline.

Assessments of character education mentioned in these interviews included evaluations that occur mostly through teacher observation (Meidl, 2008).

Character Development

Berkowitz, Bier, and MacAuley (2017) assert that the point of character education is to develop people that are a force for good. This assertion illustrates that the concept of character education and character education assessments is rooted in the preconception that people's character can be developed. Because character education is about development, it is important that character education programs follow the pathways of development in the expression of character attributes (Callina & Lerner, 2017). In that same way, assessments also need to take into account a developmental model of character (Card, 2017).

Research on character development is somewhat new and changing. For many years, it was believed that character was something that could not be changed, especially after adolescence (Clement & Bollinger, 2015). However, current research is trying to understand how character does change as it is influenced by nature and nurture (Clement & Bollinger, 2015). Some believe that moral development is a development of a cognitive understanding of values that is demonstrated through actions (Kavathatzopoulos, 1991). This concept was introduced by Kohlberg in 1958. However, Piaget also followed a similar thought process by

tying character development to his cognitive stages of development. While the stages of moral development constructed by Kohlberg and Piaget differ in their goals and methods, they are both centered on how changes in cognition affect moral reasoning and character (Kavathatzopoulos, 1991). More recent students of Kohlberg have developed a more multifaceted version of his moral development concept. In this model, there are four parts to moral development: moral sensitivity, moral judgment, moral motivation, and moral character (Curzer, 2014). This model seeks to explain how some people develop an understanding of the definition of right actions but not the habit of choosing to act on that definition. The reason for this phenomenon is that it is possible to develop different aspects of morality at different rates. In this model, the development of character is related to but is not the same as the development of understanding morals. Instead, character develops as people become sensitive to moral decisions, know the right decision to make, are motivated to act on that decision, and then habitually make that right decision. People develop different aspects of these four parts of morality through the influence of both nature and nurture over the course of their lives (Curzer, 2014).

While the Neo-Kohlbergian theory of moral development is just one idea of moral and character development, it agrees with Farrington (2017) that character development of any kind is not a linear process where children move from one level of character to the next over time. Farrington (2017) notes that character development is not just about behavioral development or neurological development. It is a phenomenological process which means that it is related to the individual experience of consciousness. Farrington (2017) argues that children develop character through what they experience and then the meaning that they make of those experiences. This means that while students may grow in their knowledge about character with age, their true character growth is related more to their experiences than to their intellectual knowledge.

Farrington (2017) argues that this approach follows a modern understanding of learning where learning is defined as obtaining knowledge through experience. Learning experiences include both opportunities for students to act and to reflect. By the process of acting and reflecting, children create habits and neurological pathways that become their character (Farrington, 2017).

Character Education Assessments

Definition of Assessment

According to Brookhart & Nitko (2019), assessment in education includes any information that is collected from students for the purpose of making decisions about instruction, policy, programs, or anything else relating to education. This can include any kind of test or non-test that is used to collect information (Brookhart & Nitko, 2019). Assessment is a key resource in the instruction of any subject because it gives teachers valuable information about the effectiveness of the instruction. For the purpose of this literature review and research, the terms assessment and measurement will be used synonymously.

Lack of Assessment

Park and Peterson (2009) explained the link between assessment and the educational value placed on character education by saying, “One measures what one values, and one values what one measures” (p. 4). Some believe that since schools and teachers value character education, they need to demonstrate the value of character education by assessing growth in character in students (Davidson, 2014). Just as in other subjects, evidence from assessments will drive an understanding of the effectiveness of character education programs (Beatty, 2017). Even though many research projects have correlated high levels of character to success in colleges and careers, there is still little assessment happening in character education programs to measure their effectiveness (Davidson, 2014). For example, in one study of a school in

Indonesia, researchers Luh Gd Rahayu Budiarta, Luh Putu Artini, I Ketut Seken, and Made Hery Santosa (2018) studied how consistent teachers' planning, teaching, and assessments were in character education. There were eighteen character traits mandated by the government to be taught in Indonesian schools. Through a series of interviews and observations, it was found that the teachers planned for ten of the traits, explicitly taught fourteen traits, but only assessed four of those traits. In the English classes in the school, none of the traits were assessed in any way (Budiarta, Artini, Seken, & Santosa, 2018). This is just one example of planning for character education and even teaching it but not making time in the class schedule to assess it in order to understand instructional effectiveness. A lack of assessment can often demonstrate a lack of importance placed on character education (McKown, 2017).

While the above example is from Indonesia, it offers a glimpse into what the United States Department of Education cited as a problem in 2001. While "character education" is included in many schools, it is not consistently implemented or held to accountability standards. Also, an understanding of what measures can be used to show the effectiveness of character education is unknown to many teachers (United States Department of Education, 2001).

Types of Assessments

Three major aspects of students' character can be measured through assessments. These are the cognitive, affective, and behavioral domains of character (National Center for Educational Evaluation and Regional Assistance, 2009). The cognitive aspects involve what students know about character and their ability to reason about moral issues. Affective aspects include the attitudes and emotions that students hold toward character. The behavioral aspect is how students behave to demonstrate positive or negative character traits (National Center for Educational Evaluation and Regional Assistance, 2009).

In 2016, Alexander detailed the different methods of assessments in character education and offered suggestions to improve these assessments. There are two different methods of assessment and measurement in character education. One method is quantitative and is focused on specific actions and attitudes. The other is qualitative and is focused on the reasons behind what people do. Qualitative approaches to assessment in character education stem from beliefs that students must have moral reasoning to live in a liberal democracy. Four ways to improve moral education measurement were suggested. These were the following: researching the cultural contexts of a moral education program, defining a good that considers human freedom, analyzing the definition of good, and creating multiple methods of evaluation for knowing whether the good has been reached (Alexander, 2016).

Quantitative assessments. As Alexander (2016) said, assessment in character education centers on knowing whether the “good” of character has been reached. While there are both qualitative and quantitative assessments, focus in research has been on quantitative assessments such as statistical surveys, suspension rates, and scales.

An example of quantitative assessments of character is a using surveys paired with academic results. For example, in a study conducted in schools in urban Chicago, Bavarian and colleagues (2013) used student and teacher self-reporting on the character quality of motivation. Students and teachers answered questions about student motivation through numerical scales. Also included in this study were measures of students’ academic outcomes. After the implementation of a character education program, motivation scores and academic achievement both increased (Bavarian, 2013). This is just one example how the effectiveness of character education is measured through quantitative assessments. Many other studies include measures of

growth in academic outcomes as arguments for the power of character education (Bavarian et al. 2013, Snyder 2007, Jeynes 2019, Skaggs & Bodenhorn, 2006).

Others assess the quality of character education through the decline in negative behaviors or suspension rates. If character education's goal is to teach children how to live out a certain set of positive values, then children who are learning those positive values should also be acting in positive ways. In theory, a character education program that is fulfilling its goals should see a decline in negative behavior and suspensions. One study that followed this line of thought was done in the state of Hawaii by Snyder and colleagues (2009). They used multiple quantitative indicators of the success of a popular character education program known as *Positive Action*. These indicators included suspension, absenteeism, grade retention, and academic achievement scores. By choosing some elementary schools as intervention schools and others as control schools, the effectiveness of Positive Action to change these measures was studied. In the intervention schools, after four years of Positive Action, suspension, absenteeism, and grade retention rates all decreased. At the end of the journal article, the authors argue that character education should be included in schools because it decreases negative behaviors (Snyder, 2009). This study is one example of how quantitative measures of suspension and absenteeism rates can be used to argue for the effectiveness of a character education program.

Another method for measuring character with quantitative means was designed for use in the military and is called the Character Assessment Rating Scale. This scale was developed by Hendrix, Luedtke, and Barlow (2004). In this scale, there is a list of character-related behaviors that people use to self-assess and then assess others. The scale uses numbers 1-8 where one represents never and eight represents always. The individuals being assessed scale themselves first. Next, a superior, a peer, and a subordinate rate the individual. Average scores on the scale

increased with the education level of the individual (Hendrix, Luedtke, & Barlow, 2004). While this scale was developed for adults, it is an interesting example of the type of quantitative assessments available to be developed in character education. Scales such as CARS are the most-used assessment model for character education. In a survey of different assessments used to measure the effectiveness of one-hundred character education programs, ninety-seven used scales or subscales to demonstrate effectiveness (National Center for Educational Evaluation and Regional Assistance, 2009).

Qualitative assessments. Other assessments of character and character education programs are qualitative. Qualitative measurements do not focus on objective numbers but on subjective perceptions. Some methods of qualitative measurements are observations and self-reporting open-ended surveys.

Observation is a popular method for assessment of character education in individual classrooms. Teachers observe the actions and words of their students and then make conclusions about the students' character. Assessments in character education lessons include actions that students do to demonstrate their character while the teacher observes the action. For example, in one study done by Robertson-Kraft and Auston (2015), teachers planned for and then implemented lessons in character. In 85% of the lessons, the teachers included some way for the students to apply their learning of a character trait through practice. In one lesson, the teacher instructed the students to practice grit by pushing themselves to do a plank for as long as they could (Robertson-Kraft & Austin, 2015). In this situation, the teacher would use observation of how students respond to the task to assess their learning about the character trait of grit.

Another way to use qualitative means to assess character is through self-generated methods such as self-reports, personal goal setting, and journaling. These self-generated

methods are argued for by Park and Peterson (2009) in their evaluation of the teaching of character strengths in school. In this version of measuring character, students use a descriptive assessment to identify character strengths and areas of growth. Next, the student can identify a character trait that they want to grow in and create goals to practice it. Every so often, they journal about their progress to self-report their growth (Park & Peterson, 2009). For example, a particular student might want to grow in honesty. That student could make a goal that related to his or her individual life such as, "I will be honest about how I am feeling with at least one friend every day." Then, the student would be given an opportunity to journal about his or her struggles or growths in honesty. While this does not provide a teacher with a grade or a score, it allows students to take charge of their own character growth and gives educators information about the students' perceptions of their own character. Park and Peterson (2009) suggest that this method of self-reporting is a more individualized method of character education that is more effective and meaningful to students.

Combined methods. In some situations, both quantitative and qualitative methods are used to assess character education. Lovat, Clement, Dally, and Toomey (2011) combined quantitative and qualitative data to produce results that found how values education affected student achievement, behavior, and school ambiance in schools in Australia. The main resource for this was pre- and post-surveys given to students, parents, and school personnel. The interviews resulted in statistically significant results in teachers' positive perception of students' growth in positive behavior. This theme continued in qualitative interviews with both teachers and students (Lovat et al., 2011). This is one example of a method of assessment that uses both qualitative and quantitative measurements.

Nature of Effective Assessments

In the *Journal of Research in Character Education*, McKown (2017) detailed some of the essential qualities of effective character education assessments. He argued that discussion on methods for measuring learning in character education must include clarity about the definition of character, assessments created for real use by real teachers, openness about the reasons for character assessment, and methods for assessment that fit the character being assessed. Researchers, teachers, schools, and others must define what they mean by character before attempting to assess it. Also, assessments used to measure character have to be evaluated for whether they work in classrooms and schools in real life. Teachers and schools must be clear that the purpose of character education assessments is to inform instruction. Finally, just like with assessments for other subjects, the method of assessment should match what is being assessed. For example, some aspects of character are better measured with surveys, while others are better measured by observation or peer reports (McKown, 2017).

Researchers must be able to develop accurate measurements of character. Card (2017) argues that good measurements of character must include the three fundamental psychometric properties. These properties are reliability, validity, and measurement equivalence. If an assessment is reliable, it will continue to give the same results with the same conditions. The validity of an assessment refers to whether it accurately measures what it is intended to measure. Measurement equivalence says that an assessment should perform the same no matter the context. Making character measurements that are reliable, valid, and have measurement equivalence can be difficult to create. This is because aspects of character are often not defined, are measured in diverse contexts, and are changeable. However, measurements of character

should still strive for high levels of reliability, validity, and measurement equivalence (Card, 2017).

Dangers of Assessment

Even with a recent push to improve assessments, character can be difficult to measure because of how complex of a concept it is. Also, there are many different definitions of character. In an attempt to identify agreed-upon positive character traits, Park and Peterson (2009) listed twenty-four different character strengths. However, they acknowledge that from this list, there are more dimensions to character. Character is demonstrated through actions, feelings, words, attitudes, and more. Also, character traits are expressed in individuals to differing degrees across a continuum. This multi-layered nature of character can make it difficult to measure. Therefore, Park and Peterson propose that character must be measured by self-reports which take into account the individuality of each student (Park & Peterson, 2009). However, as Clement and Bollinger (2015) point out, self-reports tend to be affected by the reporters' false perceptions of self or desires to answer what is socially appropriate. Clement, Bollinger, Park, and Peterson hit on an important consideration that forces educators to ask questions about the plausibility of authentic assessment with such a complicated issue.

Another possible danger of assessment is described by Pryiomka (2018) in her analysis of a study measuring the character quality of grit. In the analyzed study, grit was correlated to higher levels of academic achievement using a quantitative self-reporting scale. However, Pryiomka questions the use of the scientific method to identify character traits that correlate to any results. She argues that scientific measurement systems for character most often become tools to sort populations into the desirable and undesirable. The character traits valued are the ones which align with the middle or working-class students. This value system creates a system

where, in order for students to have “good character,” they must align with the values created by systems of population stratification. She also speaks into the implausibility of validity checks on quantitative measures of character. Surveys and polls are open to the participants’ own personal understanding. This means that when researchers write a question about respect, their idea of respect could be extremely different from the concept of respect in the mind of the participant answering the question (Pryiomka, 2018).

Method

For this case study, the research method combined elements of grounded theory and guidance from procedures advocated by Stakes (1995) and Merriam (2009) to conduct an instrumental embedded case study. The research began with an initial question. Though research on assessments of character and character education programs is growing, it is important to consider what teachers think about assessing character. While the best way to answer this question might be a large-scale qualitative research design, that research project was beyond the limits of an undergraduate research project. Instead, a more straight-forward case study could offer insights into how a singular case could address the research question. While this case cannot be generalized to all schools and teachers, it does offer valuable knowledge to the field by offering unique qualitative research. Three main kinds of case study research are defined by Stake (1995): intrinsic, instrumental, and collective. Intrinsic case studies are chosen because of the uniqueness of a certain case. The importance is not the overall theme but the case itself. An instrumental case study is focused on an idea and uses one case to offer a window into that idea. A collective case study is an instrumental case study where there are multiple different cases (Stake, 1995). This research study was an instrumental case study because it began with research questions and used the case of teachers from one school in an urban setting to offer

answers. Also, this is an embedded case study. This means that there is one case with multiple parts. The case is teachers in an urban school setting who teach grades K-5. However, there are six individual teachers in the case. The elements of grounded theory included in this study were related to the data analysis. Instead of different themes being imposed on the data, the themes described later arose from the data through a process of coding. The outline of this paper follows a pattern described by Merriam (2009) in her description of how to present findings from a case study. Lastly, the entire research construct assumes an epistemology of constructivism.

Constructivism means that knowledge obtained, interpreted, and understood is constructed by both the giver and the receiver (Stake, 1995). In other words, constructivism assumes that those interviewed were developing their knowledge and perceptions of the issue moment by moment throughout the interviews. Interviewees could have been influenced by any number of things, including the environment, the time of day, and the researcher's questions and appearance. Because of all these factors, something like human perception is best described qualitatively instead of analyzed quantitatively. Quantitative methods often do not take into account the power and also fluidity of perceptions (Stake, 1995). For all of these reasons, the method of the study was a qualitative instrumental embedded case study.

Research Questions

RQ 1: How do teachers perceive assessments in character education?

RQ 2: How does the grade level that teachers teach impact their perceptions of assessments in character education?

Case Selection and Description

The case for this study emerged from a series of events related to the researcher, location, timing, and nature of the research question. The researcher is an Elementary Education major

and was seeking to understand answers to the research question in the elementary context. Also, the researcher attends university in an urban area. After receiving approval from the school district, the researcher identified a school for the case study. The case study school is a public elementary school in an urban community in the mid-Atlantic region. The school has an average student population of 600. In the year before the research data was collected, about 15% of the students were chronically absent. In 2016-2017, 95% of the students qualified for free or reduced meals (“Virginia School Quality Profiles,” n.d.). This school allowed the researcher to contact all the K-5 teachers in general education classrooms at the school. This excluded the teacher that the researcher was working with for a required course practicum, avoiding any conflict of interest. This case included six different teachers. There were three kindergarten teachers, one first grade teacher, one third grade teacher, and one fifth grade teacher. The participants ranged in years of experience from a first-year teacher to veteran teachers. The ages and ethnicities of the teachers were also varied. However, all the teachers were female. The teachers were given pseudonyms which are as follows: Talia, Jill, Liz, Sue, Eliza, and Anna.

Design and Procedure

As mentioned above, the researcher began by gaining approval from the school district and the school principals. Then the researcher set up interview times through email with the different participants. The interviews followed a semi-structured design.

The questions included in the original interview protocol were:

1. Can you tell me about your role in character education as a teacher here?
2. How do you share information about students’ character with parents or guardians?
 - a. How do you share with the school administration?
 - b. How do you share with the community?

3. What would you think if the school board created a policy requiring teachers to track students' progress in character education programs? What would you want them to know about character education in the classroom?
4. Does the grade level of your students impact how their character is measured?
 - a. If no: Should grade level impact how students' character is measured?
 - i. If yes: Why?
 - ii. If no: Why not?
 - b. If yes: How does the grade level impact how character is measured?
5. Is there anything else you would like to share with me about measuring the character of students in school?

The researcher began with these questions. However, the researcher also chose different questions based on the participant's responses throughout the interview. Also, following the model of an open-ended interview protocol, the researcher added questions to the interviews after transcribing some of the first interviews. After each interview, the researcher transcribed the interviews. The researcher performed data analysis using methods derived from grounded theory for the purpose of an embedded case study. The researcher began by coding the research. From this analysis, the researcher developed themes. From there, the researcher sent the transcribed interviews to two different coders. These coders both independently developed five themes that agreed with themes that the researcher developed. After discussion, the researcher and coders came to an agreement on two additional themes. These themes were categorized into four major themes with sub-themes.

Data Analysis

The data revealed three themes about how the teachers in this case study perceived assessments or measurements of character in character education. Also, it revealed a theme that linked teachers' perceptions of assessments to different grade levels. The three themes about teachers' perceptions of assessments in character education are as follows: teachers equate character assessments to behavioral assessments, teachers believe that assessments could be a helpful tool, and teachers see some difficulties with assessing character. The theme about grade level is that the grade level that the teachers teach impacts their views of types of assessments and how the development of students impacts assessment.

Similarity of Character Assessments and Behavior Assessments

Throughout the interviews, teachers of all grades equated assessments of character to assessments of behavior. This focus on the behavioral domain of character in assessment is consistent with the research. According to the National Center for Educational Evaluation and Regional Assistance in 2009, assessments that measured behavioral outcomes were the most common type of assessments used to research the effectiveness of character education programs. This equation of character with behavior was evidenced in some of the teacher's use of language. For example, Sue, a first-grade teacher, used the words character and behavior interchangeably. When asked about a student whose character improved while he or she was in the teacher's class, Sue talked about how she supported one student so he knew:

Sue: that I'm here to help him change his character, change his behavior.

For some teachers, this focus on behavior stemmed from a belief that students' behaviors are the best evidence of their character. This is what Talia, a kindergarten teacher, explained when asked about her understanding of the relation of behavior and character:

Talia: I feel like your behavior is often described through your character... you can measure character through students' behavior.

She later explained that she felt like she had focused on behavior throughout the interview because she was seeing students' character through their behaviors.

Because character and behavior were considered either the same or connected, many teachers answered questions about measurements of character with by referencing types of behavior management systems. One of these systems was Class Dojo which is a points system where points are rewarded or rescinded based on behaviors that are often linked to character traits. For example, if a student picks up another student's pencil, that student can receive a point for "Kindness." However, if the student hits another student, that student could lose a kindness point. Other systems included sending folders home to parents with faces based on the students' behaviors throughout the day, giving or taking away free time, and rewarding schoolwide awards for positive behaviors. All of these systems are based on students' behavior and were mentioned by the teachers as part of character education or character assessment.

Another character assessment used in this case study school is character grades. In this system, the school identified a list of character qualities it desired to instill in its students. Every grading period, the students receive a "grade" for each character quality. These character grades are described with an "S" for satisfactory, "I" for improving, and "NY" for not yet. When asked about how these grades are assigned, the teachers spoke to their personal observations of students' behaviors. Eliza mentioned some of the questions she asks about students to herself as she assigns their character grades. For a character quality of responsibility, she asked, "Do you have what you need on a day to day basis? Are you bringing what you need to school?" Liz, a kindergarten teacher, utilized a behavioral tracking sheet for some of her students who were

struggling with behavior and planned to use that tracking to decide what character grade to give the child. These methods of assigning character grades emphasized the behavioral focus of many of the teachers.

While there are three domains of character that can be measured (as identified by the National Center for Educational Evaluation and Regional Assistance in 2009), the teachers in this case study focused on the behavioral domain of character. This focus on behavior was accompanied by an emphasis on the use of assessments that measure behavior.

Character Assessments as a Helpful Tool

Some of the teachers in the study believed that it would be helpful to have more uniform assessments for character education. These teachers spoke of the need for greater consistency from grade to grade in the expectations for students' character. They worried that while they had strong expectations, other teachers might not have those same expectations. Also, they mentioned that while the school uses a uniform system of character grades, there is not a uniform method for assigning the grades. Some teachers based their grades on students' growth in character over time, and others based their grades on the amounts of behavioral incidents for specific students. Anna, a fifth-grade teacher, wanted a system where the standard for a character grade was based on the frequency of positive or negative behaviors. She also mentioned that objective standards for character education would help her to set goals based on measurable growth in character:

Anna: I think that it [character] being something measurable for all teachers, I think that would help because I think even for the teachers, we have to be able to set goals and the kids have to be able to set goals.

Sue, a first-grade teacher, also responded positively to the idea of more a standardized character education program:

Sue: I think that when a school or school district has a behavior management or character education system in place that is school-wide or district-wide, the kids... they stick to it better.

Talia, a kindergarten teacher, wished that there was a more defined and consistent character education plan in the school. Also, she responded positively to the idea of tracking progress in character education:

Talia: It's a good idea... I think that it would help to hold the kids accountable. Maybe they could do their own tracking of their behavior or their own tracking of character.

With these comments, some teachers in the study affirmed a desire to have more consistency and organization in assessments in character education. They agreed that it would be helpful for students' understanding and teachers' planning and implementation.

Difficulty of Assessments

However, on the other side, teachers acknowledge that character is difficult to assess or measure. There were four different sub-themes that all describe different issues with assessing character. These are as follows: theoretical probability, practicality, the influence of students' home lives, and weakness of suspension as an assessment.

Theoretical plausibility. A major reason in the teachers' minds for the difficulty of assessing character was the subjective and situational nature of character. Eliza questioned how it would be possible to have an explicit way to track students' progress in character education. Her reasoning behind this was this:

Eliza: I guess I just don't see how to assess it; you know. Unless you are just in those moments day to day when they have those opportunities to be organically kind or helpful or respectful and you just happen to see that... Because you can't really put them into a situation, like a created situation to see how they would respond...

Jill, a kindergarten teacher, mentioned the subjective nature of assessing character. She explained that this subjectivity stems from the fact that different teachers have different perceptions of character traits and acceptable behaviors. Sue, a first-grade teacher, said that it was tough to measure character because character education should be about growth, not hitting a certain standard:

Sue: If you start grading [character] on certain standards or things like that measuring based on certain standards, then you couldn't be more objective about it. And I think that character is... more about growth. And I think if you were to be measuring, I think it should be measured on growth because no child who... has horrible character is going to all of a sudden wake up and be the best person in the whole wide world.

These are just three examples where teachers wonder how it would be possible to measure character since character can be difficult to define and pinpoint in a standardized way.

Practicality. Some teachers were concerned that assessments are not practical within the school day. They mentioned that they were already too busy with the current expectations, and they felt that they do not have the time to add another assessment to their already full workload. This was especially true for the older grades, where the pressure of state-standardized testing is intense. Both the third-grade teacher and a teacher talking about her experiences in fifth grade mentioned that the pressure of state-standardized testing squeezes out time for character education. The latter said this:

Liz: If you were to ask an upper-grade teacher to track this [behavior for character grades] for every child, she would never be able to teach... If I'm tracking behavior so much that I can't teach or my instruction is inhibited, their success rate goes down because they are not learning what they need to learn so now they don't pass the [state-standardized tests].

Talia, a kindergarten teacher, expressed concerns that there would be too much planning and time involved in tracking progress in character education. Jill, another kindergarten teacher, explained that she believes that state-standardized tests have changed education; teachers are under such a time crunch that they cannot focus on character as much as in the past. She said that while more standards or assessments for character might be a good idea, teachers are already trying to support the growth of their students in academic subjects and academic assessments. Jill even pointed to the fact that state-standardized tests are what determine the future of a teacher and even a school:

Jill: [Teachers] are focusing more on the classes that have tests that go with them because tests determine if your school is accredited, which determines if your school stays open, which determines if you have a job... That stress is there for a lot of teachers, and that's very sad but very true.

From these quotes, the evidence is strong that the teachers believed that assessments for character education would not be practical within the school day. With the pressure of students' academic and standardized tests, they feel that character education is pushed to the periphery in both time and importance.

Students' home lives. Measurements of character are highly influenced by home environments. The teachers felt that many of their students are not seeing good character at

home or that the students' parents do not care as much about character education as other content areas. The teachers believed that it is harder to teach character to these students and therefore, measurements are more difficult to use. While the interviewed teachers are attempting to communicate with parents about the character measurements, they were concerned that parents do not care about measurements. Some teachers believed that the parents of some of their students do not want to hear anything from school. Other teachers mentioned that some parents only want to see their child's academic grades, not character grades:

Liz: Parents aren't reading those comments [about character grades] ... They don't read that part of it. They look at the grade section, especially the core curriculum, and that's it. So once again you're at that tug-of-war with home.

Anna: I don't really care for those grades very much... I have to be honest with parents, but parents don't really look at that. I think they look more at the grades than the bottom half. I had a parent say, "What are you getting an I in?" And they went down the list to see, and it was one of the character education things and her next response was, "That doesn't matter. That's not a grade."

While the teachers might have perceived assessments a certain way in isolation, the beliefs about the parents' perceptions of the assessment can change how teachers think about an assessment.

Another idea from the interviews was that assessments in character education could also be difficult because the homes of the students impact their behaviors sometimes more than the school's character programs. Liz spoke about how assessments of the effectiveness of a character education program need to take into consideration the home environments of the

students. She even mentioned the possibility of not being able to find a character education program that can break the cycle of bad character traits being reinforced at home:

Liz: So, effectiveness... I don't know. How do you say whether or not [a program is effective] ... because we're fighting a battle that we're going up against the culture of the society and the house? It's not about what our building is doing; you can't say it's about what our character program is... So maybe it's not about the school's character development programs, maybe there is no development program that can overcome what is happening at home and the culture of their house. I don't know.

These teachers brought up a difficulty of character assessments that could be missed. They suggested that the effectiveness of character education programs is not the only thing influencing the measured behaviors of students. Students' actions are also influenced by home lives and the parents' perceptions of character education. This can add another complicated dimension to character education assessments.

Suspension rates. Many of the teachers believed that suspension rates are not an effective measure of character education. When asked about how effective suspension rates are as an assessment of the character of students in a school, teachers spoke to the complicated picture that suspension rates quantify. Liz explained that there has been a shift in educational thought about suspension that has prompted schools to try to keep students in class as much as possible. This means that some students with some serious behaviors might not be suspended. Eliza felt that suspension rates did not reflect the truth about students in a school:

Eliza: When it comes to suspension rates, a lot of higher-ups don't suspend because they don't want a higher rate. So, it's like yes, we have a low suspension rate, but the building

is one fire... Yes, the suspension rate is low, but is the classroom environment healthy?

And sometimes, it's not.

While some programs use suspension rates to assess the effectiveness of character education, the teachers in this case study talked about the difficulties of using suspension rates for this purpose.

Grade Level Taught and Perceptions of Character Assessments

There was a difference in how teachers from upper and lower grades perceived character education assessment. The main difference was that teachers in lower grades focused on behavior reinforcements like Class Dojo to measure character; however, when teachers who have taught or are teaching older grades talked about character, they focused on feeling restricted in the assessment strategies by the students' loss of external motivation and home lives.

In kindergarten, teachers focused on basic character qualities and skills such as kindness or honesty. These were reinforced through behavior management strategies like Class Dojo or behavior reports in take-home folders. However, the teachers of upper grades in elementary (3-5) spoke less about behavior reinforcements and more about just modeling traits. The different focus of grade-level teachers could be because almost all of the teachers believed that the way that character is taught and measured should be differentiated for different grade levels. The main difference noted was the rigor of expectations for the students. Eliza, a third-grade teacher, agreed that grade level impacts character education measurements because older students should have a more in-depth knowledge of character. Anna, a fifth-grade teacher, believed that older students should be held to stricter behavioral standards:

Anna: I think the upper grades should be held to a higher standard and accountability...

When I taught second, ...if they were mean, I just gave them an improving, if they are constantly mean no matter what... I just gave them an improving.

Talia, a kindergarten teacher, emphasized the developmental changes between age groups. When asked if the grade level of the student impacts how their character is measured, she said:

Talia: Maybe not how it is measured, but how far it is developed. I mean it is hard for them [kindergarteners] to talk about things that are abstract, but I think I would measure it the same.

Liz, talking about the difference in age groups, said,

Liz: We in lower grades, we absolutely use behavior as teachable moments... At this point [younger grades] what I say to them has an impact. In older grades, what you show them has more of an impact. Because all they hear are words. Don't tell me what to do. Show me how to act.

This idea that cognitive knowledge growth over time should impact character is consistent with Piaget's theory of moral development where different cognitive stages relate to different moral stages (Kavathatzopoulos, 1991). It also indicates that teachers assume older students have been taught more about character than younger students.

However, Eliza also spoke to a disconnect between the older students' knowledge and their actions:

Eliza: So, I think to a certain extent, like yes, age matters, but at the same time, we can't assume that the older they are, the more they know. Because with our school, what we

are seeing is the older they are, they struggle more even though they should know better in theory.

Other teachers also saw a disconnect in the theoretical knowledge of older students and their actions. Liz expressed her frustration with this section of her interview:

Liz: In a functioning society, by the age of eight or nine, you should feel like, “Well I know what I do or say is going to affect them [my neighbor]” and have that much of a knowledge. And instead we have gotten to a place where they have the knowledge, but they are, “I don’t care. It’s all about me so I don’t care.” So how do you teach them empathy and sympathy? ...They’ve become numb, that age kid. You could promise them the moon- don’t care not going to do it... So how do you instill [character] when they’re already so defeated? And I don’t know the answer to that.

In summary, the difference among teachers of lower and upper grade levels was the types of assessments on which the teachers focused. They also answered a question about the differences in assessments for grade levels. While some said that assessments should be differentiated since older students should meet higher expectations than younger students, others countered that while older students might know more, they do not always act on that knowledge.

Researcher Bias

The researcher’s bias could have impacted this case study in a few different areas. For one, before creating the interview questions or analyzing the data, the researcher did a preliminary literature review. During this review, the researcher began to form opinions about assessments in character education. Also, the researcher had hypotheses about how the teachers would respond to different questions. To mitigate bias, the researcher requested that her thesis

chair review her interview questions. Also, as mentioned in the method section, a process for validity was followed during data analysis by the use of multiple coders.

Implications

While a single case study such as this one cannot be generalized to all cases, it still has implications for educational researchers, policy-makers, administrations, and teachers. If character education is mandated and mandated education requires accountability, then assessments for character education are the logical conclusion. However, as evidenced by the research as well as these teachers' voices, character is a complex issue. Educational research on assessments for character education cannot merely focus on quantitative measures that leave out the voices of those who are most involved, the teachers. If assessments in character education are to become a research-based practice, all types of research should examine it. Just from this small case, it was evident that teachers have some expectations for and qualms about assessments of character. More qualitative case studies should be done in schools of all grade levels and in diverse environments to bring further light to the breadth of teachers' perceptions. Policy-makers and administrations must begin to catch a glimpse of how teachers perceive the plausibility and practicality of assessments. While many teachers have not participated in formal research studies, they are highly qualified to speak into the complexities of character education by personal experience. Before assessments are developed or mandated, they should be filtered through teachers' experience. Lastly, individual teachers can begin to think about how they perceive character and measure it in their classrooms. They can take into account their understanding of what character is, the practicality of integrating different types of assessments, students' home lives, and the developmental stage their students find themselves in.

Conclusion

While character education is a concept that many teachers are favorable to and state legislations even mandate, it is not treated in a similar way to other types of education in schools. For one, it is not assessed in similar ways or with similar consistency. In an educational climate that values assessments, character education has received more research recently about possibilities for assessments. However, in this research, the voices of teachers have been left out. Whether this is because of the novelty of research in assessments in character education or because of the lack of value placed on qualitative research, the perceptions of teachers in character education assessment is an entirely new arm of research. This research emphasizes the importance of the teachers' voice in making educational decisions that are considered research-based. It also asks questions about difficulties in character education as a concept to which some teachers alluded. The answer to the struggles to assess character and growth in character might go beyond even insufficient research to something much deeper for which research, whether quantitative or qualitative, might not have the answer.

References

- Alexander, H. A. (2016). Assessing virtue: Measurement in moral education at home and abroad. *Ethics and Education, 11*(3), 310-325. doi:10.1080/17449642.2016.1240385
- Bavarian, N., Lewis, K. M., DuBois, D. L., Acock, A., Vuchinich, S., Silverthorn, N., . . . Flay, B. R. (2013). Using social-emotional and character development to improve academic outcomes: A matched-pair, cluster-randomized controlled trial in low-income, urban schools. *Journal of School Health, 83*(11), 771-779.
- Beatty, A. (2017). Afterword: Reflections on the special issue. *Journal of Character Education, 13*(2), 65-66.
- Benninga, J., Berkowitz, M., Kuehn, P., & Smith, K. (2006). Character and academics: What good schools do. *Phi Delta Kappan, 87*(6), 448–452.
- Berkowitz, M.W. (2012). Understanding effective character education. CSEE Connection.
- Berkowitz, M. W., Bier, M. C., & MacAuley, B. (2017). Effective features and practices that support character development. *Journal of Character Education, 13*(1), 33-51.
- Brookhart, S., & Nitko, A. (2019). *Educational assessment of students*. London: Pearson Education.
- Card, N. A. (2017). Methodological issues in measuring the development of character. *Journal of Character Education, 13*(2), 29-45.
- Character Education Legislation (n.d.). In Character.org.
- Clement, S., & Bollinger, R. (2016). Perspectives on character virtue development. *Research in Human Development, 13*(2), 174-181. doi:10.1080/15427609.2016.1172445
- Curzer, H. J. (2014). Tweaking the four-component model. *Journal of Moral Education, 43*(1), 104–123.

- Davidson, M. (2014). A character education research perspective for the 21st century. *Journal of Research in Character Education, 10*(1), 77–83.
- Farrington, C. A. (2017). Illuminating the mechanisms of character development. *Journal of Character Education, 13*(1), 53–65.
- Hendrix, W. H., Luedtke, C. J., & Barlow, C. B. (2004). Multimethod approach for measuring changes in character. *Journal of Research in Character Education, 2*(1), 59-80.
- Hunter, J.D. (2001). *The death of character: Moral education in an age without good or evil*. New York, NY: Basic Books.
- Jeynes, W. H. (2019). A meta-analysis on the relationship between character education and student achievement and behavioral outcomes. *Education and Urban Society, 51*(1), 33-71.
- Kavathatzopoulos, I. (1991). Kohlberg and Piaget: Differences and similarities. *Journal of Moral Education, 20*(1), 47.
- Leming, J. S., & Yendol-Hoppey, D. (2004). Experiencing character education: Student and teacher voices. *Journal of Research in Character Education, 2*(1), 1-17.
- Callina, K. S., & Lerner, R. M. (2017). On the importance of coherence in the study of character development. *Journal of Character Education, 13*(1), 17-25.
- Lovat, T., Clement, N., Dally, K., & Toomey, R. (2011). The impact of values education on school ambience and academic diligence. *International Journal of Educational Research, 50*(3), 166-170. doi:10.1016/j.ijer.2011.07.008
- Lovat, T., & Dally, K. (2018). Testing and measuring the impact of character education on the learning environment and its outcomes. *Journal of Character Education, 14*(2), 1-22.

- Luh Gd Rahayu, B., Luh Putu, A., I Ketut, S., & Made Hery, S. (2018). How consistent is teachers' planning, implementation, and assessment in character education? *SHS Web of Conferences*, 42, 61. doi:10.1051/shsconf/20184200061
- McKown, C. (2017). Promises and perils of assessing character and social and emotional learning. *Journal of Research in Character Education*, 13(2), 47-52.
- Meidl, C. (2008). *Teachers' beliefs and attitudes toward approaches to character education* (Order No. 3336090). Available from Education Database. (304496510).
- Merriam, S. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: John Wiley & Sons.
- Milson, A. J. (2003). Teacher's sense of efficacy for the formation of students' character. *Journal of Research in Character Education*, 1(2), 89-106.
- National Center for Educational Evaluation and Regional Assistance. (2009). *Survey of Outcomes Measurement in Research on Character Education Programs* (NCEE 2009006).
- Park, N. & Peterson, C. (2009). Character strengths: Research and practice. *Journal of College & Character*, 10(4), 1–10.
- Pryiomka, K. (2018). Quantitative dogmatism in character assessment and its implications for education: A case study of grit. *Social and Personality Psychology Compass*, 12(5), e12388-n/a.
- Robertson-Kraft, C., & Austin, K. (2015). The character of achievement: An analysis of teachers' instructional practices for character education. *Journal of Character Education*, 11(2), 109.
- Sikes, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.

- Skaggs, G., & Bodenhorn, N. (2006). Relationships between implementing character education, student behavior, and student achievement. *Journal of Advanced Academics, 18*(1), 82-114.
- Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., Beets, M., & Li, K. (2009). Impact of a social-emotional and character development program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: A matched-pair, cluster-randomized, controlled trial. *Journal of Research on Educational Effectiveness, 3*(1), 26-55.
- Sorrells-Blackmon, F. M. (2017). *A comparative study of middle school teacher character education efficacy beliefs in Northeast Georgia* (Doctoral Dissertation). Retrieved from Digital Commons.
- Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage Publications.
- United States Department of Education. (2001). No child left behind act of 2001. In <http://www.ed.gov/policy/elsec/leg/esea02/pg69.html>.
- Virginia School Quality Profiles. (n.d.). In <http://schoolquality.virginia.gov/>.
- Yin, R. (2009). *Case study research: Design and methods*. Thousand Oaks, CA: Sage Publications.