

2005

Developing and testing a "responsibility" questionnaire for 8th graders

Melissa S. Ward
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Ward, Melissa S., "Developing and testing a "responsibility" questionnaire for 8th graders" (2005).
Graduate Theses, Dissertations, and Problem Reports. 734.
<https://researchrepository.wvu.edu/etd/734>

This Thesis is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Thesis in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Thesis has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

Developing and Testing a “Responsibility” Questionnaire for 8th Graders

Melissa S. Ward

Thesis submitted to the
College of Human Resources and Education
at West Virginia University
in partial fulfillment of the requirements for the degree of

Master of Arts
In
Educational Psychology

Daniel E. Hursh, Ph.D., Chair
Paul E. Chapman, Ph.D.
Floyd L. Stead, Ph.D.

Department of Advanced Educational Studies

Morgantown, WV
2005

Keywords: Values Education, Ethical Instruction, Content Validity, Delphi Technique,
Internal Consistency, Test Development

Copyright 2005 Melissa S. Ward

Abstract

Developing and Testing a “Responsibility” Questionnaire for 8th Graders

Melissa S. Ward

The purpose of this study was to develop and test a questionnaire that can assess students’ attitudes and behaviors as they relate to the trait “responsibility.” Content validity was obtained using a modified version of the delphi process. Several experts in the area of measurement design or character education contributed suggestions. The questionnaire was revised based on these suggestions. The questionnaire was completed by a small group of eighth graders. The reliability of the questionnaire was estimated using a test of internal consistency. The questionnaire was written by this author, with the guidance and assistance of professionals who have experience and expertise in the area of test construction. The test items were based on the definition of responsibility provided by the Character Counts organization. The questionnaire includes twenty Likert-type items with eleven of these measuring perceptions of behavior and nine measuring perceptions of attitude.

Internal consistency tests showed Cronbach’s Alpha to be .834 for the entire questionnaire, .752 for the questions measuring behavior and .637 for the questions measuring attitude. Internal consistency tests should be repeated with additional pilot groups to obtain further evidence that the questionnaire is reliable. Future refinement of the questionnaire should incorporate other methods of establishing validity and reliability – for example, a determination of whether or not the results of the questionnaire are predictive of certain behavioral problems.

Table of Contents

Abstract.....	ii
List of Tables	iv
Introduction.....	1
Evaluating Character Education	4
Character Education in West Virginia	5
Review of the Literature: Introduction	6
Review of the Literature	7
Method	11
Participants and Setting.....	11
Measurement.....	11
Procedures.....	12
Data Analysis and Research Design	13
Results.....	13
Validity	13
Reliability.....	14
Discussion.....	14
Validity	14
Reliability.....	15
References.....	17
Appendix A.....	19
Appendix B.....	22
Appendix C.....	25
Appendix D.....	28

List of Tables

Table 1	29
Table 2	30
Table 3	31
Table 4	32
Table 5	33
Table 6	34
Table 7	35
Table 8	36
Table 9	37

Introduction

Perhaps the elders of every age are more or less troubled about the ‘morals’ of the oncoming generation...[b]ut at certain periods of especially hurried transition this uneasiness becomes more than customarily acute. (Peters, 1933, p. 214)

The quote above seems as if it could have been written today, considering the ongoing technological revolution and the pervasive world violence of our time. These words were, in fact, written in 1933 by Charles C. Peters in his introduction to a series of character education studies. Even then, there was apparently much concern about the character of America’s youth – and about whether or not character could indeed be taught in school.

However, character education is not solely an American issue; it is a theme that has surfaced again and again throughout the histories of many nations and cultures. An early example of character education appears in the teachings of Socrates. Socrates used arguments of logic to challenge his students’ beliefs and attitudes. His rhetoric, subtle as it was, was a powerful tool (Cole, 1950). In ancient India, students who sought to study with the best teachers were rejected if they were found lacking in character, and students could not “graduate” until they were found to be intellectually *and* morally competent (Evans, 2000). A more current example is modern England, where a review of recent research indicates that character education has received as much “explicit official attention” in England as it has here (Halstead & Taylor, 2000, p. 169).

In America, character education has always been present in one form or another, in part because our democratic society expects its citizens to possess certain character traits such as responsibility and self-discipline (Lavaroni, 1999). One example of an early connection in the U.S. between character education and the school is the McGuffey Reader. The McGuffey Readers, developed by William McGuffey in the 19th Century, taught students the importance of virtues such as honesty, hard work, kindness, and courage (Lavaroni, 1999). The McGuffey Readers integrated moralistic lessons and selections from the Bible with non-religious lessons covering traditional content areas such as spelling and science (Scully, 1967).

In the 20th Century, new ideas about moral development and education were presented in the research and writings of Jean Piaget and Lawrence Kohlberg. Jean Piaget was a Swiss biologist who, later in his life, became interested in psychology. Piaget developed a theory of intellectual development which states that individuals move forward through stages of intellectual development; regression to previous stages does not occur. Piaget was interested in social development as well, and through his research he deduced that cognition and affect work in synchrony as children move through the stages of social development. Piaget’s research in this area led to his interest in moral judgment. Although Piaget proposed that moral development is also composed of stages, he never worked out the details of these stages.

Piaget became influential in the United States in the late 1950s; it was during this period that Lawrence Kohlberg, an American psychologist, was drawn to Piaget’s work. Kohlberg eventually created his own theory of moral development, inspired and based on Piaget’s ideas. The official title of his theory is the *cognitive-developmental theory of*

moralization – the title emphasizes that there is a cognitive basis to moral development. The assumption underlying Kohlberg’s theory is that moral judgment is a direct function of cognition. Moral development and cognitive development are not perfectly parallel to each other, but logic does play a major role in moral development. Kohlberg’s model of moral development contains six stages, with each stage having specific characteristics based on the moral *reasoning* used by the person in that stage. According to Kohlberg’s theory, an individual may sit on the borderline between stages, but will not regress backwards. Kohlberg’s theory remains popular, although many criticisms have been proffered throughout the years – for example, there is little evidence that Stage 6 truly exists (although Kohlberg found that some philosophers reasoned at this level); only boys were used in Kohlberg’s studies; and he changed methods throughout his research, which was longitudinal by necessity (Reimer, Paolitto & Hersh, 1983).

Kohlberg felt that his theory could contribute to moral education in the school in two ways. The first was through incorporation of moral instruction into the curriculum; the second was through a restructuring of the school environment. Through these methods children were to be exposed to moral reasoning that was one stage higher than their own so that they could be coaxed up into that stage. For teachers, this meant recognizing their own level of moral development; learning how to handle moral dialogue; and recognizing stages of moral development within students (Reimer, Paolitto, & Hersh, 1983).

In the 1960’s, a form of character education called values clarification gained popularity as a reflection of personalism – the belief that one’s personality is unique and should be kept safe from the influence of others. Although this form of character education differed significantly from past forms – due to its insistence that students choose their own values – its former popularity demonstrates that the schools continued to play an important role in character formation (Lavaroni, 1999).

Today, the persistent popularity of character education can be seen in the multitude of books, articles, websites, and organizations that focus on every conceivable aspect of character education. Some controversy has sprung from this widespread interest, as scholars and practitioners disagree about which form of character education is most effective. One topic of this debate is the “word of the month” form of character education, which has received criticism due to its implied tenet that there are specific traits that “should” be taught to students.

Given the various forms character education has taken over time, one might wonder *what exactly* qualifies as character education. There is no one answer to this question. One might begin to define character education by defining character. One definition of character is offered by Peters: “Character consists of an aggregate of habits, attitudes, and functioning philosophies of life” (Peters, 1933, p. 215) This definition is comprehensive because it emphasizes both behavior and attitude. However, there are certainly as many definitions of character (if not more) as there are forms of character education.

Indeed, as shown above, “character” or “moral” education can take a variety of forms. Cline and Feldmesser identify four primary types of moral education (1983). The developmental model was created by Lawrence Kohlberg who, inspired by the work of Jean Piaget (see above), proposed that individuals move through sequential stages of moral development just as they move through stages of cognitive development.

Kohlberg's theory ultimately inspired an instructional program in which students are presented with moral dilemmas and are encouraged to consider each dilemma from the viewpoint of an individual who is in a moral stage one step above that of the student. Kohlberg felt that the whole school should be structured to represent a higher stage of moral development than that of its students to facilitate student movement into a higher stage of moral reasoning.

Values clarification was created by Louis E. Raths, who believed that values should be freely chosen by the individual and consistently acted upon. The instructional method based on Raths' theories requires that teachers challenge students to clarify their values; teachers are never to suggest that one value is better than another. Raths' emphasis on values as a personal choice is based on the belief that schools should not indoctrinate students nor should they be proponents of any particular value.

The actionist approach is a "learning by doing" model of moral education. Action learning is a combination of direct experience in the community and related instruction in the classroom. Direct experience is not restricted to community service but may include other community interactions such as internships or submersion in another culture. The ideal action learning environment allows students to make important decisions; requires them to perform challenging tasks; and teaches them how to cooperate with others. Action learning is designed to reduce or remove barriers between students and the community and to give students opportunities to contribute to their communities – this may in turn reduce apathy, cynicism, and even delinquency. Action learning connects to Kohlberg's moral development theory because significant experiences within the community, combined with time for reflection, may spur moral deliberation and moral growth. Action learning may also boost academic achievement; because students must apply what they learn in the classroom, they may be more motivated to learn (Hedin & Schneider, 1978).

The rationalist approach stresses the importance of using rational thought to solve problems and holds up legal institutions as examples of sophisticated moral structures. The related instructional method centers on how the legal system works and attempts to foster positive attitudes towards the legal system. Sometimes referred to as citizenship education, this approach emphasizes rational thinking, community involvement, and the democratic process (Kentucky State Department of Education, 1977).

Each of these models has its own strengths and flaws, and there seems to be no evidence to suggest that any one of these models is more effective than the others. A fifth model of character education not identified by Cline and Feldmesser is the "word of the month" approach. Schools that use this model introduce one character trait to students each month; teachers are to integrate the trait into their normal instructional activities, and the administration may develop school-wide activities that focus on individual traits. Variations of this model may include a "word of the week" approach or a temporary emphasis on a specific trait such as respect. Again, there seems to be no evidence that "word of the month" models or any variations thereof are any more effective than the others listed above.

Of course, one might look beyond all of the models of character education that have gained and lost popularity throughout history and ask, "Why teach character at all? How do we justify this?" One common response is that teachers cannot help but impart morals to their students through their choice of curricular materials, instructional

methods, and more informal interactions with students. Some would argue that the responsible thing for teachers to do is to be *aware* that they are doing this and to make *conscious* decisions about how they go about imparting morals to their students.

Evaluating Character Education

In determining why character should be taught, school personnel should determine what it is that they want from a character education program. Is the goal for students to *apply* what they learn about character just as they might apply what they learn in math and history and other subject areas? Are students developing and refining certain character traits or simply memorizing the definitions of these traits? These questions represent one reason that character education programs should be evaluated – evaluation can indicate whether students are applying what they learn or simply memorizing words. Another reason is that character education is a controversial area of study, in particular because experts disagree about what is the most effective form of character education. For example, some experts believe that “word of the month” programs are ineffective because they focus on individual character traits instead of focusing on moral reasoning in general (Kohn, 1997). “Word of the month” programs, then, need to be investigated because if they are indeed ineffective, perhaps schools that use such programs should implement some changes. One might hope that through research, the controversy over what methods are most appropriate might be diminished. One additional reason for evaluating character education is that the public has been disillusioned in the past by innovative programs that didn’t seem to work (Cline & Feldmesser, 1983). Since character education is required by law in West Virginia (and many other states), it is important for the public to have evidence that this mandated program is producing positive results.

A final reason we must evaluate character education is to learn *how* to evaluate character education. The concept of morality can seem vague and intangible, and efforts to capture the effectiveness of a program may sometimes prove to be ill-conceived and futile. Only through repeated and persistent efforts can we learn how to effectively evaluate character education programs.

Before one even begins an evaluation of a program, the question must be posed: Can a character education program *be evaluated* at all? Some, like Robin Barrow, would argue that there is no definitive answer. Argues Barrow, “[S]ociological or psychological inquiry that is scientific will inevitably fail to tell us anything reliable about human beings: it will assert universal laws where there are none” (Barrow, 2000, p. 316). The implication for educational research, Barrow suggests, is that seemingly straightforward questions (e.g., does character education work?) do not in fact have an answer. The reason is that human beings differ so much from one another that there cannot be one educational method that will work for all.

I believe that it is worthwhile to attempt an evaluation of character education programs, but I do recognize that any individual attempting such an evaluation faces certain obstacles. For example, there appears to be a limited selection of appropriate measurement tools for students. For a “word of the month” type program, the ideal attitudinal or behavioral survey for students would be comprised of questions or statements that correlated directly with the *definitions* of character traits written or chosen by those providing character education. This ideal survey would be carefully written and

tested for reliability and validity. Unfortunately, there seem to be few such ideal surveys available.

Another obstacle to evaluating character education is the apparent lack of structure in character education programs. Often teachers are permitted to integrate character education as they see fit, and so the treatment differs from teacher to teacher. This creates some problems:

- Treatment cannot be clearly defined
- Treatment will not be the same for all students
- Scores may be misleadingly high for students that have had 'more' character education than their peers

An additional obstacle to evaluating character education is the definitional issue. For a "word of the month" type program, official definitions of character traits may be available to all teachers; however, it is not guaranteed that all teachers will consult these definitions as they decide how to incorporate these traits into their lessons. This means that even if appropriate measures were found or created, there would be no guarantee that they would be valid measures of what the students were actually learning.

I decided to focus on the lack of appropriate measurement tools by attempting to create and test my own instrument. Before one can attempt to evaluate a character education program, an appropriate instrument must be found or developed. A poorly developed or inappropriate instrument will ultimately tell the researcher very little about the effectiveness of the program. In the case of "word of the month" type programs, an appropriate instrument would not measure character in general, but would measure changes in students that related directly to the specific character traits included in the program.

Character Education in West Virginia

West Virginia's House Bill 2208 amended state code to include the following: "(a) The state board shall establish a comprehensive approach to integrate character education into all aspects of school culture, school functions and existing curriculum. (b) The state board shall require all public schools. . . from preschool to grade twelve to develop and integrate components of character development into their existing curriculum" (H.B. 2208, 2001, ¶ 1-2). West Virginia's resultant project mission for its character education program is stated as such: "All students in West Virginia will exhibit core character traits that are necessary in order to become productive and successful citizens" (West Virginia DOE Abstract, 2001, ¶ 2).

In north central West Virginia, as well as in some other areas in West Virginia, the Character Counts program – a "word of the month" model – is often used to fulfill this mission. According to a newsletter prepared by one county's board of education, this model uses "a common language to describe and teach character" that "substantially increases the effectiveness of character-building programs by exposing young people to consistent and pervasive messages" (Fike and Henline, 2002, p. 3). The decision to use the Character Counts program is further explained in another newsletter, in which character education is defined as "the web of instruction and experiences that schools, families, and the community provide through teaching and. . . impact the behavior of children and youth in [the areas of] respect, responsibility, caring, trustworthiness, fairness, and citizenship." (Vetere, 2002, p. 5).

Because the Character Counts program is popular not only in West Virginia but throughout the United States, I chose to design an instrument that could be used to evaluate one of the traits advocated by this program. Specifically, I decided to focus on the trait “responsibility.” Ultimately, my goal was to design and test an instrument that would measure students’ behaviors and attitudes as they related to the trait “responsibility.” I wanted to create an instrument that was reliable and valid, and to make this instrument available to others so that they could use it and/or test it further.

The intended purpose of this instrument is to assess student attitude as well as behavior. Many studies of character education have focused on changes in student behavior as measured by teacher observation or disciplinary reports. It seems that few studies actually incorporate student-generated perceptions of attitudes and behaviors. Behaviors can change for reasons other than attitudinal change towards specific values (i.e., tougher disciplinary policies, threats from parents, etc.). The results of a more comprehensive instrument will reflect a change in student thinking regarding specific values and the behaviors that are the observable evidence of those values.

Review of the Literature: Introduction

Before beginning a survey of the available studies on character education and discussing the instruments used in these studies, one should note that there is much disagreement among experts about what character education should entail. According to Alfie Kohn, character education can refer to a general attempt to enhance children’s character, but the term is generally used to refer to what he calls the “kit” version of character education, where individual traits are presented in neat packages. Kohn claims that with this form of character education, indoctrination is the preferred method of instruction. That is, children are seen as “blank slates” who must be *told* what is right instead of learning for themselves. He notes that positive reinforcement is a popular component of such programs, even though the studies he has reviewed have shown that positive reinforcement doesn’t work with character education – in fact, it has an adverse effect. “The lesson a child learns from Skinnerian tactics is that the point of being good is to get rewards,” he says (Kohn, 1997, p. 182).

Kohn goes on to argue that the political and economic realities of children must be addressed by character education programs – children are not born “bad,” and many of their behaviors are reinforced by personal circumstances that character education (in the narrow sense of the word) may not address. He also laments that character education does not address social environment, which also plays a part in the way people behave (for example, residents of a small town may be more likely to help a stranger in need than residents of a large city). This concept of social environment is transferable to the school environment. We cannot just change the students, Kohn claims. We must change the learning environment as well.

One additional problem with programs that focus on specific values, according to Kohn, is that educators are often not careful about the values they choose and how they define and illustrate these values. He warns that educators should not equate any value with obedience to what teachers and administrators say (i.e., saying that getting in a straight line is a form of respect). Vague terms such as “respect” can be slippery.

Wynne provides a contrasting viewpoint (1995). He supports programs such as “word of the month” programs that focus on specific character traits and he believes that

there *can* be consensus about what values to teach. “[W]ho . . . is against teaching honesty, diligence, or kindness?” he asks, although he does not address the issue of how such traits should be defined or what specific behaviors are associated with these traits (Wynne, 1995, p. 151). Wynne is critical of programs in which separate values are not specified; he actually claims that an increase in youth crime and other negative behaviors correlates with the past popularity of “values clarification” programs. Wynne also is an advocate of prohibition and punishment, saying that the human “dark side” is an unfortunate reality that must be dealt with at school – this is counter to Kohn’s implication that humans are inherently good. Wynne seem to mistakenly associate “values clarification” programs with other programs, such as those proposed by Kohlberg, in which students are challenged to progress to a higher level of moral development. There are programs that writers such as Kohn might support, in which students are not taught specific values but instead are expected to refine their belief systems. Educators who use such programs are encouraged to coax students away from negative belief systems; students are not, in fact, encouraged to choose their own values.

There are numerous other models and theories relating to character education in the general sense of the term. The presence of such multitudinous models, with all the contrasting viewpoints of each model, necessitates that a variety of models be evaluated by researchers so that those that are clearly ineffective can be eliminated by educators. The following review of literature focuses on those programs that emphasize specific traits. As will be seen, many of these studies are flawed due to choices of instrumentation.

Review of the Literature

Many studies of character education programs seem to rely on perceptions of change; in other words, surveys are conducted only after character education has been implemented, upon which students, staff members, or parents are asked whether or not they feel the program has made a difference. For example, teachers in several school districts in Utah conducted their own action research projects to determine the effectiveness of their character education programs. One teacher administered a survey to students and staff asking them to “rank statements about each trait [that was being taught through character education] and how the students in our school related to the trait” (Utah State Office of Education, 2000, p. 16). There is no indication that any such survey was administered *before* the character education program was implemented in that school and it seems that students were required to make assumptions about their peers. Although students may be more honest about other students’ behaviors than they are about their own, it would be impossible for them to accurately assess the actual attitudes of their peers. Other projects focused on issues such as how character education was being implemented and how students and staff felt about a character education program that had been eliminated (Utah State Office of Education, 2000). Gathering such information might be useful, but this information does not indicate the relative effectiveness of the character education programs in those schools.

Another example is an evaluation of a program called the Heartwood program that was implemented in elementary schools in a school district near Pittsburgh (Lanke, Wood & Gettleman, 1991). The primary feature of this program was reinforcement of positive character traits through exposure to global literature (where the stories’

characters demonstrated these positive traits). The project was carefully planned and preceded by a survey administered to superintendents and curriculum coordinators that asked whether or not they felt there was a need for a character education program in their district. The authors of the study determined that the program was effective; however, they based this conclusion, in part, on statements made by teachers after the program was implemented, with no corresponding interviews or surveys conducted beforehand. Students were asked to tell stories that illustrated character traits, and the researchers documented the percentage of students who correctly identified and illustrated specific traits. This storytelling activity was conducted during the program implementation, with no corresponding activity occurring before program implementation. The only pre/post component of the study involved student ability to correctly identify the locations of certain countries in which some of the stories were set. The effectiveness of the program from the students' point of view is unknown; all that was learned from the students was whether or not they could identify character traits.

Another study was conducted in three schools in a suburban metropolitan area (Duer, Parisi & Valintis, 2002). The program was designed to increase respectful and responsible behavior and to decrease disciplinary office referrals. Activities included the teaching of conflict-resolution skills; teacher modeling; empathy building; group discussions with targeted students; and positive reinforcement. The researchers did compare disciplinary referrals before and after the program was implemented; such a comparison can yield useful information about changes in student behavior. However, other measurement choices were problematic. The students and staff were given surveys beforehand that measured their perceptions of student, teacher and administrator behavior; they were asked to evaluate others' behaviors without any consideration of their own behaviors. After the program had been implemented, students were administered a survey in which they were asked whether or not they had noticed a change in others' behaviors; these post surveys were not parallel to the surveys administered before the program was implemented. Again, students were not asked to consider their own behaviors, and perceptions were not measured at all.

In one study, a researcher sought to evaluate the effectiveness of a "word of the month" program in a small private elementary school (Grades K – 8) in San Francisco (Lavaroni, 1999). The chosen character traits were "respect" and "responsibility." No rewards other than praise were used due to conflicting studies about the results of the use of extrinsic rewards in character education programs. Teachers were given the word of the month and asked to incorporate it into discussions and assignments. However, the minimal involvement allowed was a posted sign, so technically the teachers did not have to even speak the words "respect" and "responsibility." It seems that the principal did ensure that the students completed one writing assignment a month based on the given word.

Eight teachers completed surveys. Six of the teachers thought the program was effective and felt that it 'brought an awareness' of what the words "respect" and "responsibility" meant. They also noted that the students at the very least used the words more often. Two of the teachers weren't sure, with one teacher noting that there were no clear goals for the program and so its effectiveness was difficult to assess. One teacher questioned whether or not improved behaviors could be attributed solely to the new program. Three of the teachers wondered about the long-term effect. Unfortunately the

students were never asked their opinions, and the fact that they used the words more often does not in any way indicate that there was an increase in positive behavior. Additional flaws that may have influenced the effectiveness of the evaluation is that the terms “respect” and “responsibility” were never clearly defined, which explains why some teachers may have seen the goals as being inchoate. Without these definitions, there is no assurance that the results were valid. Another problem was that the treatment was not consistent from teacher to teacher, and there seemed to be no attempt to measure student change according to the actual treatment used. The evaluative data consisted only of teacher observation and “anecdotal information” (Lavaroni, 1999, p.8).

A comprehensive evaluation of a character education program in Utah could have been more well-developed. More than 65 schools in Salt Lake City and surrounding communities participated in this program through a grant, and all were evaluated to some extent (Johns, 2001). Some of these schools were involved in the program for up to four years. Each school was permitted to design its own program, and the decision of which values were to be emphasized varied from school to school. Many schools chose to use a “word of the month” type program; teachers were asked to integrate the values as they saw fit. Service-learning projects were also developed to reflect the schools’ dedication to the application of the values being taught. Other programs such as peer mediation were implemented as well.

The evaluator took great care to use several different measurement methods in order to assess several aspects of the program. These included pre- and post-testing of students, parents, and staff on the Character Development Survey (developed by the evaluator); post testing of students on a measure entitled Knowledge of Character Education Concepts (also developed by the evaluator); the analysis of pre- and post-data from student disciplinary records; and an implementation survey administered to teachers at the end of the grant to document professional development and the instructional strategies used. The results indicated that 62% - 80% of students (figures come from two separate administrations) correctly defined a set list of values. Some of the values may not have been taught to all students, possibly causing the lower scores of the initial administration. This, then, was not necessarily a weakness in the instrument, except that the instrument did in some cases not reflect the treatment. The results also indicated that students felt that expectations regarding behavior were made clear almost all of the time. In response to questions about changes in student attitudes and behavior, 73% of the site facilitators (one for each school) reported positive changes, which included an increase in positive behaviors as well as a decrease in behavioral problems (the low number of 73% might suggest that schools that had already been implementing character education before the provision of a grant would not yield noticeable changes in student behavior – however, there were no previous studies with which to confirm this). Fifteen schools reported decreases in negative student behavior as evidenced by disciplinary records.

The evaluator admits, in her report, that the evaluation procedures included many problematic elements. One of these was that the schools had varying experience with character education, so that results from different schools could not be compared to one another. Lack of baseline data was also a problem, since the evaluation measures were developed after the project began. These problems underline the importance of developing evaluative procedures before a program is implemented, establishing baseline data, and then evaluating after a set period of time instead of waiting, say, four years.

Another measurement-related problem was that not all measures had “pre” counterparts, and on some measures post-tests – which one might argue are more important than pre-tests -- were optional. Yet another measurement-related issue was that disciplinary records were collected haphazardly, with some schools providing post-records only, and still others being able to provide none due to flaws in their data gathering systems. The many measurement flaws of this study caused the overall results to be less valid and reliable than they could have been, although the use of several measurement tools counterbalances this somewhat by providing many sources of data.

The Character Counts organization – which developed the program used by many schools in West Virginia and across the nation – compiled the results of assessments of its program (Josephson Institute, 2003). All of the studies in this compilation concluded that the program yielded positive results; however, most of the studies consisted only of comparisons of before and after disciplinary records. One exception is an extensive five-year study in South Dakota that included thousands of elementary and secondary school students and teachers. The study showed that crime rates and drug use declined between 1998 and 2000. In addition, the number of students claiming to have engaged in behaviors such as shoplifting dropped. Negative school behaviors such as cheating and skipping class also decreased. There were two important findings in relation to this study: the first was that the program had the most dramatic effects on students in grades one through six. The second was that the more exposure the students had to the program, the better they behaved. It appears that the student survey used for this study asked questions about attitudes and behavioral changes, but these questions did not explicitly address the Six Pillars (character traits) chosen by the developers of Character Counts. However, it is possible that assumptions about specific traits could be made based on the responses to questions about specific behaviors listed in the survey.

In less comprehensive studies completed in Virginia and Louisiana, teachers reported positive behavioral changes in relation to the Six Pillars. The Virginia study – which included a survey given to elementary school teachers – reported statistically significant changes. This survey asked teachers to judge 24 kinds of student behavior – four behaviors for each of the Six Pillars. The creators of this survey did, then, attempt to relate the survey items to the specific traits taught, but it is unfortunate that there was no corresponding survey for students. The Louisiana study also relied on teachers’ perceptions of change; no surveys were given to the students. One study that actually asked for students’ perceptions of change occurred in 1999 at an elementary school in West Des Moines, IA. Students here generally responded positively when asked about changes in their behavior, changes in other students’ behaviors, and whether or not learning about the character traits was important. The questions on this survey were primarily “yes” or “no” questions that did not refer to specific character traits.

Overall, the available research seems to show certain pervasive measurement-related weaknesses in those studies that address “word of the month” type programs. Those who are evaluating such programs must ask and answer more questions – in particular, questions about students’ perceptions – so that educators will have a more lucid understanding of the relative effectiveness of these kinds of programs. Rather than be discouraged by weaknesses in this area, researchers can and should design additional studies that address unanswered questions and attempt to correct previous weaknesses.

In doing so, they must concentrate on developing or finding reliable and valid instruments and using these instruments appropriately.

Method

Participants and Setting

The participants in this study were middle school students. Middle school students were chosen because it is at this age that many behavioral problems surface. This occurs, in part, because of differences in the way adolescents make decisions; adolescents often find themselves dealing simultaneously with a new-found autonomy and intense emotional changes, both of which can affect their ability to make good decisions (Kipke, 1999). Specifically, the participants were eighth graders at a middle school in north central West Virginia. Eighth graders were specifically chosen because these students are in a transitional phase – they are preparing for high school, where they will be given more freedom and will be expected to show more responsibility. The instrument, however, could easily be adapted and tested for seventh graders, ninth graders, and possibly other age groups as well. There were approximately 150 eighth graders at the middle school; unfortunately, due to the low return rate of the parental consent forms, only ten were available to take the survey.

This middle school was chosen due to the socioeconomic diversity of the student population. The school included grades six through eight and had a total of 492 students. The average class size was 23.1 with a 15 to 1 pupil/teacher ratio.

Measurement

The measurement tool used was a questionnaire for students (see Appendix C). This questionnaire was designed to measure students' perceptions of their behaviors and attitudes as they pertained to the trait "responsibility."

I wrote the questionnaire with the guidance and assistance of professionals who have experience and expertise in the area of survey construction or character education. The suggestions of these experts, to the extent that it was feasible to do so, were incorporated into the final survey. Print resources were consulted for additional guidance during the construction of the questionnaire.

The questionnaire items were based on the definition of "responsibility" that is provided by the Character Counts organization: "Do what you are supposed to do; persevere; keep on trying; always do your best; use self-control; be self-disciplined; think before you act – consider the consequences; be accountable for your choices" (The Josephson Institute, 2004, *The Six Pillars of Character*, ¶ 4).

The definition was broken down into seven domains. The questionnaire contained at least three Likert-type items for each domain; some of these items asked for students' perceptions of attitude, while others asked for students' perceptions of behavior. The statements on the questionnaire were divided into two sections, with the first section containing the statements about student behavior and the second section containing the statements about student attitudes.

Each statement was followed by four possible responses, ranging from negative to positive. When writing each statement, careful attention was given to that statement's relationship to the definition of responsibility provided by the Character Counts

organization. To facilitate the creation of this relationship, it was necessary to have operational definitions for some of the terms used in the definition of responsibility provided by Character Counts (see Appendix D). I believed that if all of the statements correlated with the definition of responsibility, then the questionnaire would be a valid measure of responsibility as it is defined by Character Counts.

Throughout the development of this questionnaire, it was also necessary to ensure that the reading level was appropriate for eighth graders. Word choice, item construction, and written directions were simplified as much as possible, and an item asking students to identify words or statements that were confusing to them was added at the end of the questionnaire.

Procedures

To obtain content validity, I used the delphi process. The delphi process can be defined as a process that “assesses group opinions by utilizing representative experts who work toward consensus on an issue” (Eason, 1992, p. 2). The delphi process involves multiple stages. Experts must be chosen and contacted; these experts must be given an opportunity to respond to the researcher’s questions and documents; these responses should be combined and, when feasible, integrated into the researcher’s work; and the collective response and integration of responses should be submitted to the experts for approval or additional suggestions (Eason, 1992). I identified and contacted individuals who have expertise in the area of survey construction or character education. These included professors, representatives of Character Counts and other character education organizations, and school system personnel. I asked these individuals to evaluate the questionnaire and to suggest corrections that would contribute to the content validity of my questionnaire.

My expert panel consisted of three professors (including one who helped develop Character Counts surveys for a school system in another state); the director of a character education organization; the research director of another character education organization; a teacher who has implemented Character Counts in her classroom; and a representative from the West Virginia Department of Education. Additional informal feedback was offered by a guidance counselor at the middle school where the internal consistency of the questionnaire was to be tested. I e-mailed copies of my questionnaire to all of the above-listed individuals (with the exception of the guidance counselor), and each individual e-mailed their comments and suggestions to me. I made corrections based on the feedback I received and sent the revised questionnaire back to each individual.

The questionnaire was administered on one day; preparation steps (obtaining consent, for example) took more time. Content validity (see above) was obtained over a period of several weeks.

I met with school representatives twice to explain my research. My goal was for these individuals to fully understand the purpose of my research before the questionnaire was given to students.

I needed consent from the students’ parents and from the students themselves before I could distribute questionnaires to the students. Consent was obtained prior to distribution of the student questionnaire. The questionnaire was given to students during one class period.

Data Analysis and Research Design

The design was a test of internal consistency and content validity for a questionnaire designed to measure the trait ‘responsibility.’

Table 1 displays the research questions, the data sources, and the data analyses for all research questions.

Results

Validity

The first draft of the questionnaire was designed to match the definition of responsibility provided by the Character Counts organization. This questionnaire was refined with the help of a professor at a large university who had experience with both survey construction and character education. This professor continued to offer guidance during later revisions. It was believed that using the Character Counts definition as a guide and seeking advice from an expert would help to establish content validity. To further improve content validity, a modified delphi process was used to collect suggestions from additional experts. The first draft of the questionnaire was sent by e-mail to two professors, a middle school teacher, and two individuals representing character education organizations. A guidance counselor at a middle school offered additional suggestions during a face to face meeting. Revisions were made based on the suggestions of these experts.

The draft of the questionnaire that was sent to these individuals can be seen in Appendix A. It should be noted that the above-mentioned teacher gave the questionnaire to some of her students (ages 10 –12); she then made comments based not only on her own observations, but on the problems her students encountered when they were given the questionnaire.

Table 2 shows the categories that the experts’ comments fell under and the number of individuals who made each kind of comment.

Comments primarily fit into two broad categories: wording of statements and formatting. Most of the comments addressed wording issues. Three individuals were concerned that the reading level of certain items was too difficult (for example, the phrase “attention to quality” in item 10). Others were worried that the statements were so general that the students would not know how to respond (i.e, they would not be able to envision a specific behavior). There was also some concern that the results of the questionnaire could be misleading due to the “correct” answer being too obvious or due to a response that could be perceived as “irresponsible” in some situations and “responsible” in others (for example, one might not describe hurting someone’s feelings as ‘irresponsible’ if the individual doing so is trying to be helpful).

I made revisions to my questionnaire based on these initial comments. I then sent the revised questionnaire to all of the individuals, with the exception of the guidance counselor who had offered her suggestions during an informal meeting. I also sent the revised questionnaire to a representative of the state department of education who did not have an opportunity to respond to the first version. Not all of the individuals had additional comments, but I revised the questionnaire based on the comments I did receive (see Table 3). This revised questionnaire can be seen in Appendix B.

I made additional changes to the questionnaire based on the above comments; one other change was requested by the county board of education representative who approved the study. The final version of the questionnaire can be seen in Appendix C.

In addition to using the delphi process, I included an item at the end of the questionnaire that asked students to circle any words or sentences they did not understand. None of the students circled anything, nor were there any items that were left blank by students.

Reliability

Ten students at a middle school in north central West Virginia completed the questionnaire. This pilot group was much smaller than expected; parental consent forms were handed out to approximately 150 students (nearly all of the eighth graders at the middle school), but only ten of these students returned the consent forms.

Cronbach's alpha was computed for the set of twenty questions on the questionnaire and for items deleted. The results are shown in Tables 4 - 5. Cronbach's alpha for the entire questionnaire was .834, which is higher than the preferred minimum of .70. Item deleted results show that internal consistency for all items would change very little if any items were deleted. Deletion of Item 7 or 16 would increase Cronbach's Alpha to $> .85$, suggesting that these items may be weak and should, perhaps, be rewritten.

Cronbach's Alpha was also computed separately for the eleven items measuring behavior and the nine items measuring attitude. Cronbach's Alpha for items deleted was also computed for each of these sets. The results are shown in Tables 6-9.

For the items measuring behavior, Cronbach's Alpha was somewhat lower than it is for all twenty items. Deletion of Item 7 would push Cronbach's Alpha to $> .80$, again suggesting that Item 7 should perhaps be rewritten.

For the nine items measuring attitude, Cronbach's Alpha was much lower than that of both the first set and all twenty questions. Although deletion of either Item 14 or Item 16 would increase Cronbach's Alpha, they would not bring it above the preferred minimum of .70.

Discussion

Validity

The initial version of my questionnaire contained many flaws that affected its validity. This problem was alleviated through extensive revisions based on the advice of experts in the fields of character education and survey construction.

A questionnaire as a measurement tool is useless if not carefully constructed and tested. For example, if the reading level of the questionnaire is too difficult, students may either leave items blank or choose a random response. The delphi process is certainly one effective way to improve the validity of a questionnaire. Not only does it provide the author with suggestions from multiple experts, but it alleviates the problem of the author being "too close" to the questionnaire to be able to recognize needed changes.

An item was added to the questionnaire that asked students to circle any words or sentences that they did not understand. This item was added to improve content validity by ensuring that the reading level was appropriate for the students. None of the

ten students who filled out the questionnaire circled any of the items and no student left any item blank, suggesting that the reading level was appropriate for these ten students. Further testing of the questionnaire would show whether or not reading level and vocabulary choice are appropriate for a majority of eighth graders.

Reliability

It was hoped that Cronbach's Alpha for the entire questionnaire and for each set would be no lower than .70. Cronbach's Alpha for both the entire questionnaire and the set of items measuring behavior were well above .70, with the results for the first set being lower than the results for the entire questionnaire but still indicating a high level of internal consistency for this set. Internal consistency for both the entire questionnaire and the set of items measuring behavior would improve if Item 7 – "I am able to control my temper in class when someone makes me mad" – were deleted. It would be beneficial to explore why this item is problematic before deleting or rewriting it; indeed, further testing with a larger group would confirm whether this item is problematic at all.

The results of reliability testing for the second set of items – those measuring attitude – indicate that this set has a low level of internal consistency that would not be significantly improved by the deletion of any one item. Because Cronbach's Alpha increases as the number of items increases, it is possible that internal consistency for both sets and even the entire questionnaire would improve significantly if more items were added to each set (provided that the items were appropriate for the questionnaire). Additional pilot tests should be conducted to assess whether or not this would occur. Other factors, including variability between subjects in the sample, can affect Cronbach's Alpha. Therefore, future pilot tests comprised of large, heterogeneous groups of students would perhaps provide more information about the reliability of the questionnaire. Further tests and analyses should also incorporate other methods for establishing validity and reliability. To establish stability over time, one group of students could complete the questionnaire several times within a certain time period; to establish convergent validity, the results of the questionnaire could be compared to other instruments or techniques that purport to measure responsibility; predictive validity could be established by determining whether or not scores were predictive of, say, academic performance (as one might expect, if being responsible includes completing all assignments on time, asking for help when necessary, and so on). Establishing construct validity would be tricky because the questionnaire is based on a definition of responsibility provided by Character Counts. To even begin to assess construct validity, this definition would have to be evaluated and a determination made about whether or not it accurately reflected what responsibility is. This would be difficult considering the abstract nature of the concept, but perhaps worth trying.

The number of students involved was a concern and was certainly much lower than expected. It was hoped that up to 100 students would fill out the questionnaire; the low return rate of the parental consent forms made it impossible to give the questionnaire to a large number of students and time constraints made it impossible to administer the questionnaire at other schools. There could have been at least two possible reasons for the low return rate. One might have been a lack of incentive for the students; they were not required to return the consent forms, nor were they offered any kind of reward for returning the consent forms. Another reason might have been unwillingness on the part

of the teachers to encourage students to return the forms. This could have been a result of heavy workloads – especially as the school year came to a close – with teachers giving priority to mandated instructional activities over optional assistance to graduate students at the local university.

One unexpected obstacle that complicated this study was the general lack of enthusiasm I encountered from potential participants as I narrowed down my topic and searched for a cooperating school. The sentiment that was often expressed to me by principals, teachers, and other staff members was that certain state and federal mandates – in particular, No Child Left Behind – necessitated that initiatives such as character education be “placed on the backburner.” As a result, character education programs within the schools I visited were often unstructured and “bare bones.” This lack of enthusiasm may have been compounded by my role as an “outsider.” As an individual who had neither taught nor attended any of the schools in the area, I found myself in the position of a stranger asking assistance from strangers. The concept of researcher as outsider has been experienced and studied by others. For example, two researchers who were studying the development of relationships between preservice teachers and their mentors found that the “anxieties of teachers actually affected their efforts to entice schools and teachers to be involved in the study” (Sim & McCluskey, 2003, p. 1). The researchers found that the teachers had little time for visitors or extensive data collection and were, in certain circumstances, participating only because they felt pressured to do so. The researchers acknowledged that the “mentors and preservice teachers [had] far less to gain from it than we [did]. We [were] asking participants to invest personal time. This [was] an imposition” (Sim & McCluskey, 2003, p.6). Olds and Symons (1990) noted that “the complexity of this relationship [between school personnel and researchers] has the potential to create obstacles. . . [for] those who are not ‘citizens’ of the school community” (cited in Sim & McCluskey, 2003, p. 3). It is imperative that researchers who are not a part of the “school community” carefully consider how they can approach school personnel, develop a positive relationship, and garner interest and support in the research topic (especially a topic in which the researcher is more invested than school personnel). How to go about this is beyond the scope of this study but may be worth exploring in future studies.

References

- Barrow, Robin (2000). The poverty of empirical research in moral education: beyond John Wilson. *Journal of Moral Education*, (29)3, 313-321.
- Cline, Hugh F. & Feldmesser, Robert A. (1983). Program evaluation in moral education. ERIC Document Reproduction Service No. ED241541.
- Cole, Luella (1950). *A history of education: Socrates to Montessori*. New York: Holt, Rinehart and Winston.
- Duer, Marc; Parisi, Adam; & Valintis, Mark (2002). Character education effectiveness. ERIC Document Reproduction Service No. ED471100.
- Eason, Sandra (1992). Power assessment and the Delphi process. ERIC Document Reproduction Service No. ED341732.
- Evans, Susan D. (2000). Creating moral citizens: answers from the ancient, non-western world. ERIC Document Reproduction Service No. ED456958.
- Fike, Mike & Henline, Kevin (2002). Character Counts! Coalition. *Monongalia County Schools Character Education Newsletter*, (1)2, p. 3.
- Halstead, Mark J. & Taylor, Monica J. (2000). Learning and teaching about values: a review of recent research. *Cambridge Journal of Education*, (30)2, 169-202.
- Hedin, D. & Schneider, B. (1978). Action learning in Minneapolis: a case study. In R.W. Tyler (Ed.), *From youth to constructive adult life: the role of the public school* (pp. 149-167). Berkeley, CA: McCutchan.
- Johns, Jennifer S. (2001). Utah community partnership for character education: final evaluation report, 1995-1999. ERIC Document Reproduction Service No. ED463224.
- Josephson Institute (2003). Character Counts! works: the evidence. Retrieved from www.charactercounts.org/doing/survey-reports.htm
- Kentucky State Department of Education, Frankfort (1977). Character education: some questions and answers. ERIC Document Reproduction Service No. ED159107.
- Kipke, Michele D. (Ed.). Risks and opportunities: synthesis of studies on adolescence. ERIC Document Reproduction Service No. ED438052.
- Kohn, A. (1999). A critical examination of character education. In A. Ornstein & L.S. Behar-Horenstein (Eds.), *Contemporary issues in curriculum*. Boston: Allyn & Bacon.

- Lanke, Barbara A.; Wood, Patricia K.; & Gettlement, Eleanore N. (1991). The ethics crisis – is it the school’s responsibility? ERIC Document Reproduction Service No. ED345330.
- Lavaroni, Peter (1999). The effects of a word of the month program on a small independent school. ERIC Document Reproduction Service No. ED433264.
- Martinez, Maryann (1989). The impact of goal setting techniques in effectuating self-discipline. ERIC Document Reproduction Service No. ED325730.
- Olds, R.S. & Symons, Wolford C. (1990). Recommendations for obtaining cooperation to conduct school based research. *Journal of School Health* (60)3, 96-99. Cited in Sim & McCluskey, below.
- Peters, Charles C. (1933). The potency of instruction in character education. *Journal of Educational Sociology*, (7)4, 214-223.
- Reimer, Joseph; Paolitto, Diana P., & Hersh, Richard H. (1983). *Promoting moral growth: from Piaget to Kohlberg*. New York: Longman.
- Scully, James A. (1967). *A biography of William Holmes McGuffey*. Ann Arbor, MI: University Microfilms International.
- Sim, Cheryl & McCluskey, Kerryn. Academics researching in schools – making first contact. Presented at the International Education Research Conference, November 30 to December 3, 2003, Auckland, NZ. Retrieved from www.aare.edu.au/03pap/abs03.htm#S
- Utah State Office of Education, Salt Lake City (2000). Utah character education action research projects. ERIC Document Reproduction Service No. ED463223.
- Vetere, Michael (2002). FYI. *Monongalia County Schools Character Education Newsletter*, (1)1, p. 5.
- West Virginia Department of Education (2001). Integrating character development in West Virginia [abstract]. Charleston, WV. Retrieved October 11, 2003 from www.ed.gov/programs/charactered/grantees/wv.html
- Character Education Integration. West Virginia Code 18-2-13. Retrieved from www.legis.state.wv.us
- Wynne, Edward A. (1995). Transmitting character in schools: some common questions and answers. *Clearing House*, (68)3, 151-153.

Appendix A

Version 1 of Questionnaire
Responsibility: Student Perception Survey

“Do what you are supposed to do”

1. I make my own decisions without considering what teachers expect from me.

Always Usually Sometimes Almost never

2. Sometimes I make a decision to do something at school that I know I’m not supposed to do.

Always Usually Sometimes Almost never

3. I feel that, at school, there are certain things I *must* do, whether I want to or not.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

“Persevere” and “Keep on trying”

4. If I am given an assignment by my teacher, and I do not succeed the first time, I will try again until I do succeed.

Always Usually Sometimes Never

5. I tend to give up easily when my teachers give me an assignment or task that is difficult for me.

Always Usually Sometimes Almost never

6. If I am taking a class that is hard for me, I keep trying to do well in that class even if I become frustrated.

Always Usually Sometimes Almost never

7. If I am in a class that is too hard for me, I try to switch to an easier class.

Always Usually Sometimes Almost never

“Always do your best”

8. I feel that it is important for me to do my best on homework assignments.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

9. I do my best on homework assignments, no matter how hard they may be.

Always Usually Sometimes Almost never

10. When I begin a homework assignment, my goal is to get it done without attention to quality.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

“Use self-control”

11. It is important for me to control my actions at school because it might keep me from doing something I might later regret.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

12. I feel that I have complete control of my actions when I am at school.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

13. I am able to remain calm in class even if something makes me angry.

Always Usually Sometimes Almost never

“Be self-disciplined”

14. I complete homework assignments on time.

Always Usually Sometimes Almost never

15. I follow classroom and school rules.

Always Usually Sometimes Almost never

16. I feel that my teachers are responsible for making sure that I do my work and follow the rules.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

“Think before you act – consider the consequences”

17. It is okay for me to make a decision without considering all of the possible consequences.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

18. I think about the possible results of my actions before I act.

Always Usually Sometimes Almost never

19. I react to situations without thinking about the consequences – good or bad.

Always Usually Sometimes Almost never

20. If I thought that someone would be hurt by something I did, I would not do it.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

“Be accountable for your choices”

21. I never confess to doing wrong even if I know I have.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

22. When something bad happens because of a choice I make, I tell people that it is not my fault.

Always Usually Sometimes Almost never

23. I accept any negative consequences (i.e., bad grade or in-school suspension) for the choices I make.

Strongly agree Somewhat agree Somewhat disagree Strongly disagree

**The items in this survey correspond to the definition of responsibility as provided by the Character Counts! organization:*

“Do what you are supposed to do; persevere; keep on trying; always do your best; use self-control; be self-disciplined; think before you act – consider the consequences; be accountable for your choices.”

Appendix B

Version 2 of Questionnaire
Responsibility: Student Perception Survey

Directions: Circle the category that most closely matches your behavior.

1. When I am in class, I make my own decisions without worrying about what my teachers expect from me.

Almost never Sometimes Usually Almost always

2. I do things in class that I know I'm not supposed to do, such as passing notes or copying someone else's homework.

Almost never Sometimes Usually Almost always

3. If I am given an assignment by my teacher, and I don't know how to do it, I will ask for help.

Almost never Sometimes Usually Always

4. I tend to give up easily when my teachers give me homework that is difficult for me.

Almost never Sometimes Usually Always

5. If I am taking a class that is hard for me, I keep trying to do well in that class even if I become frustrated.

Almost never Sometimes Usually Always

6. I do my best on homework assignments, no matter how hard they are.

Almost never Sometimes Usually Always

7. I am able to control my temper in class when someone makes me mad.

Almost never Sometimes Usually Always

8. I turn in my homework assignments on time.

Almost never Sometimes Usually Always

9. I follow my teachers' rules.

Almost never Sometimes Usually Always

10. I think about the possible results of my actions before I act.

Almost never Sometimes Usually Always

11. If I get in trouble at school for doing something wrong, I tell people that it is not my fault.

Almost never Sometimes Usually Always

Directions: Circle the category that most closely matches your opinion.

12. I think school is like a prison because we have to follow childish rules.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

13. I feel that it is important for me to do my best on homework assignments.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

14. When I take a test in class, my goal is to answer all of the questions. I don't care if my answers are correct.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

15. It is important for me to control my actions at school because it might keep me from doing something I will later regret.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

16. I feel that I have complete control of my actions when I am at school.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

17. I feel that my teachers are responsible for making sure that I do my work and follow the rules.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

18. It is okay for me to make a decision without considering all of the possible consequences.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

19. If I thought that someone's feelings would be hurt by something I did, I would not do it.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

20. I do not like to confess to doing something wrong even if I know I have.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

21. I do not argue when I get in trouble (for example, if I get in-school suspension for breaking a rule).

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

22. Were there some words or sentences on this survey that you did not understand? If so, please go back and circle them.

Appendix C

Version 3 of Questionnaire
Responsibility: Student Perception Survey

Directions: Circle the response that most closely matches your behavior.

1. When I am in class, I do what I want to do without worrying about what my teachers expect from me.

Almost never Sometimes Usually Almost always

2. I do things in class that I know I'm not supposed to do, such as passing notes or interrupting other people.

Almost never Sometimes Usually Almost always

3. If I am given an assignment by my teacher, and I don't know how to do it, I will ask for help.

Almost never Sometimes Usually Almost always

4. I tend to give up easily when my teachers give me homework that is difficult for me.

Almost never Sometimes Usually Almost always

5. If I am taking a class that is hard for me, I keep trying to do well in that class even if I become frustrated.

Almost never Sometimes Usually Almost always

6. I do my best on homework assignments, no matter how hard they are.

Almost never Sometimes Usually Almost always

7. I am able to control my temper in class when someone makes me mad.

Almost never Sometimes Usually Almost always

8. I turn in my homework assignments on time.

Almost never Sometimes Usually Almost always

9. I follow classroom rules.

Almost never Sometimes Usually Almost always

10. I think about the possible results of my actions before I act.

Almost never Sometimes Usually Almost always

11. If I get in trouble during class, I tell my teacher that I didn't do anything wrong even if I know I did.

Almost never Sometimes Usually Almost always

Directions: Circle the response that most closely matches your opinion.

12. I would enjoy school more if we did not have to follow so many rules.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

13. It is important for me to do my best on homework assignments.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

14. When I take a test in class, my only goal is to finish as fast as possible. I don't care if my answers are correct.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

15. It is important for me to control my actions at school because this might keep me from doing something I will later regret.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

16. I feel that I have complete control of my actions when I am at school.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

17. If my teachers didn't make me, I wouldn't follow the rules or do my work.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

18. It is okay for me to make a decision without worrying about the possible consequences.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

19. I don't care if the things I say hurt people's feelings.

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

20. I do not argue when I get in trouble (for example, if I get in-school suspension for breaking a rule).

Strongly Disagree Somewhat Disagree Somewhat Agree Strongly Agree

21. Were there some words or sentences on this survey that you did not understand? If so, please go back and circle them.

Appendix D

Operational Definitions

Perseverance – Character Counts describes a person who has perseverance: “Responsible people finish what they start, overcoming rather than surrendering to obstacles. They avoid excuses such as, ‘That’s just the way I am,’ or ‘It’s not my job,’ or ‘It was legal’” (Josephson Institute, 2002, Making Ethical Decisions).

Self-control – Restraint exercised over one’s own impulses, emotions, or desires (from www.m-w.com, Merriam-Webster’s online dictionary).

Self-discipline – “[T]he ability to maintain and alter ... goal-directed behavior with minimal outside support” (from Martinez, 1989).

Accountability – Character Counts! defines an accountable person as “[a] person [who] is not a victim and doesn’t shift blame or claim credit for the work of others. He considers the likely consequences of his behavior and associations. He recognizes the common complicity in the triumph of evil when nothing is done to stop it. He leads by example” (Josephson Institute, 2002, Making Ethical Decisions).

Table 1
Data Analysis

Research Question	Data Sources	Data Analyses
Does my instrument contain an adequate sample of the domain of content that I am trying to measure? Is the format of my instrument appropriate for my intended audience?	Expert panel	Delphi process – obtain advice from experts and make requested changes; return to experts; repeat until all experts agree (to the extent that is possible) that the instrument has achieved content validity.
	Student questionnaire	Include item at the end of the questionnaire that asks students to identify statements that they did not understand; also note whether or not there were certain statements to which students consistently did not respond.
Using a measure of internal consistency will I obtain evidence of reliability for my instrument?	Student questionnaire	Cronbach's Alpha to obtain reliability coefficient.

Table 2
Questionnaire Version 1 Comments

Type of Comment	# of individuals who made this comment
Item too general	2
'Correct' answer too obvious	2
Purpose of statement not clear due to phrasing or word choice	2
Response could be misleading	1
Reading level too high/awkward phrasing	3
Format of response options should be modified	2
Overall format of questionnaire should be modified	1
Example items in directions needed	1

Table 3
Questionnaire Version 2 Comments

Type of Comment	# of individuals who made this comment
Purpose of statement not clear due to phrasing or word choice	2
Reading level too high/awkward phrasing	1
Response could be misleading	1
Wording too harsh	1

Table 4
Internal Consistency of Questionnaire

Cronbach's Alpha	N of Items
.834	20

N=10

Table 5
Internal Consistency of Questionnaire for Items Deleted

Item	1	2	3	4	5	6	7	8	9	10
Cronbach's Alpha if deleted	.817	.817	.837	.840	.833	.813	.851	.829	.811	.804

Item	11	12	13	14	15	16	17	18	19	20
Cronbach's Alpha if deleted	.825	.809	.829	.835	.815	.853	.835	.842	.808	.822

Table 6
Internal Consistency of Behavior Set

Cronbach's Alpha	N of Items
.752	11

N=10

Table 7
Internal Consistency for Items Deleted: Behavior Set

Item	1	2	3	4	5	6	7	8	9	10	11
Cronbach's Alpha if deleted	.707	.707	.761	.772	.740	.6814	.813	.737	.677	.694	.731

Table 8

Internal Consistency of Attitude Set

Cronbach's Alpha	N of Items
.637	9

N=10

Table 9

Internal Consistency for Items Deleted: Attitude Set

Item	12	13	14	15	16	17	18	19	20
Cronbach's Alpha if Item Deleted	.517	.620	.656	.554	.671	.604	.622	.577	.624