

The Assessment of the Adult Learning and Development Student Satisfaction Scale (ALDSS Scale)

Jonathan E. Messemer, Ed.D.

Catherine A. Hansman, Ed.D.

Cleveland State University
College of Education and Human Services
CASAL Department – Julka Hall 275
2121 Euclid Avenue
Cleveland, Ohio 44115-2214
United States of America

Abstract

The purpose of this article is to discuss the development and assessment of the Adult Learning and Development Student Satisfaction Scale (ALDSS Scale). The ALDSS Scale was designed to measure student satisfaction regarding the following six factor groups: (1) curriculum, (2) learning format, (3) course materials, (4) program access, (5) faculty and instruction, and (6) faculty advising. During a three year period, the ALDSS Scale was administered to students nearing graduation from an adult learning and development master's degree program. The article addresses the rate of reliability among the six factor groups and how the reliability measurements change when specific items are deleted from the ALDSS Scale. The researchers measured the construct reliability bias among the six factor groups with respect to race.

Keywords: Adult Education, Adult Learning and Development, Reliability, Graduate Students, Higher Education, Student Satisfaction, Survey Development.

1. Introduction

We in the United States are currently living in an era whereby the common political view is to demand that federal and state departments and agencies operate under strict fiscal principles. The current political trend has left the field of higher education extremely vulnerable, with respect to funding of academic programs. Brand (1997, 2000) predicted this demand for market driven higher education more than a decade ago when he saw the signs that stemmed from the market driven policies being passed by the U.S. federal government. According to Neal (1995), "American higher education faces a common but unwelcomed dilemma – to prove its worth to an increasingly skeptical and critical public" (p. 5). When state legislatures discuss issues related to higher education, Ruppert (1995) suggests that they frequently use "terms like *accountability*, *performance*, and *productivity*" (p. 11).

The current political climate has also empowered many accreditation agencies in the United States, such as NCATE, to oversee the academic accountability process for many colleges of education. Adult education graduate programs are often assigned to university departments who may also include counseling, higher education, school administration, and supervision (Merriam & Brockett, 2007). Therefore, it is rare to see a stand-alone adult education graduate program in the United States. Since adult education graduate programs are not licensure based programs, we are often left more vulnerable to scrutiny by the state departments of education because we do not have an accrediting body. Adult education programs are often seen as less important or not taken seriously as those programs that have NCATE or CACREP as their accreditation credential. However, adult education graduate programs are not loosely run academic programs in higher education. In addition to meeting the state and university mandates, adult education professors and academic programs are governed by a well defined set of standards set forth by the Commission of Professors of Adult Education (CPAE, 2010). Because CPAE is a non-accredited body, our standards are often viewed in the public and private sectors as ad-hoc at best – I might add in much of the same manner that the nine ethical responsibilities adult educators have toward adult learners were formulated and described by Wood (1996).

The majority of the U.S. states in 1992 were experiencing budgetary problems due to the national economic recession, which according to Ruppert (1995) led to many state colleges and universities receiving less funding to run its academic programs for a period of two or more years. Between the demand for budgetary reconciliation among many U.S. universities and the lack of academic identity among many adult education graduate programs, the field has seen a number of adult education graduate programs, such as those at Boston University, Cornell University, Ohio State University, Oklahoma State University, Syracuse University, and the University of Alaska at Anchorage close since the late 1980s (Knox & Fleming, 2010).

Given the demand for more accountability on the part of U.S. colleges and universities as described by Sandmann (2010), the faculty from an Adult Learning and Development (ALD) master's degree program at a large state university in the Midwest region of the U.S. decided to develop measures to evaluate their graduate program. The ALD faculty knew that they could quickly generate the student enrollment and retention numbers for the university administrators. The ALD faculty also felt comfortable in justifying to the university administrators the level of adult learning which was occurring in the classroom. For example, all ALD students were required to complete an exit strategy component, which consisted of students completing one of the following options: ALD portfolio course, comprehensive exam, research project, or master's thesis. Regardless of which exit strategy option the students chose, the ALD faculty knew that the students would have to demonstrate the level of adult learning that they had gained through their participation in the master's degree program. The one area in which the ALD faculty did not have a mechanism for program evaluation was in measuring student satisfaction. Therefore, the researchers decided to develop the *Adult Learning and Development Student Satisfaction Scale (ALDSS Scale)*.

2. Purpose Statement

The one area in which the ALD faculty did not have a mechanism for program evaluation was in measuring *student satisfaction*. Therefore, the researchers intend to briefly discuss the development of the *Adult Learning and Development Student Satisfaction Scale (ALDSS Scale)* and the degree to which the factored groups warrant reliable measurements. The purpose of this paper is to discuss the development of the *ALDSS Scale* and to measure the reliability of the six factor groups with respect to the individual survey items. This investigation is driven by the following three research questions:

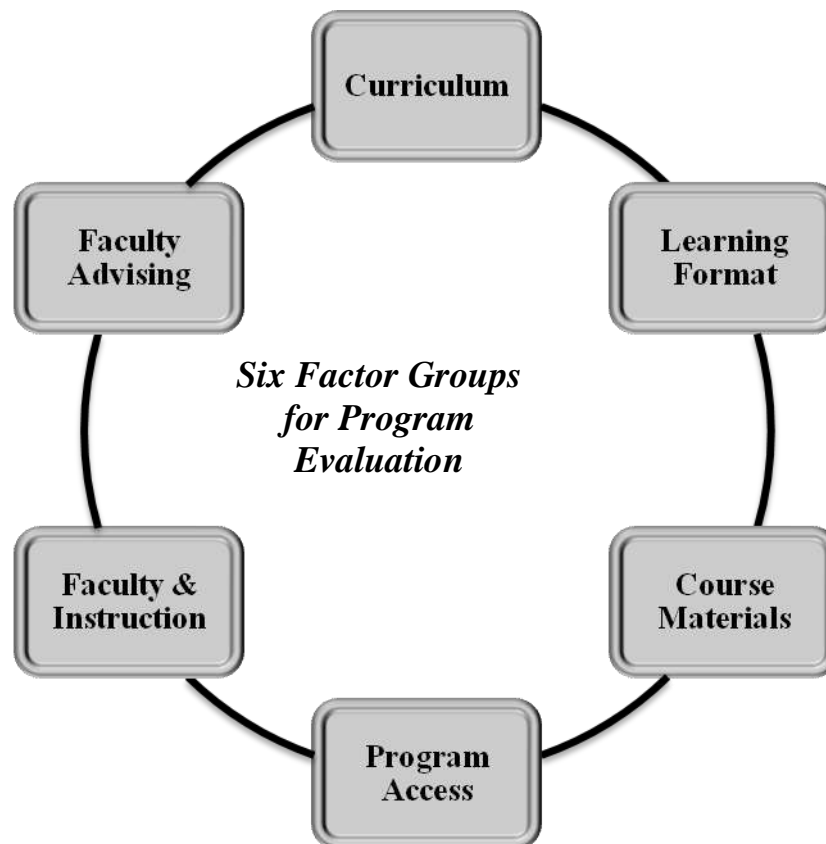
1. Does each of the six factor groups warrant an acceptable rate of reliability?
2. Does the rate of reliability increase among any of the six factor groups when specific items are removed?
3. Does each of the six factor groups warrant an acceptable rate of reliability with respect to race?

3. Literature Review

The researchers conducted an extensive literature review of previous studies measuring program satisfaction at higher education institutions. Therefore, the researchers found only four empirical studies that address some form of student satisfaction within colleges and universities (Lounsbury, Saudargas, Gibson, & Leong, 2005; Nauta, 2007; Spooren, Mortelmans, & Denekens, 2007; Zullig, Huebner, and Pun, 2009). In all four cases, the authors primarily studied undergraduate students and only the Spooren, Mortelmans, and Denekens (2007) study made note toward the psychometric values among the factor measures as it pertained to construct reliability. The researchers did not find any student satisfaction studies that pertained to the evaluation of graduate programs.

The empirical study from Spooren, Mortelmans, and Denekens (2007) sought to identify the important factors that influence the level of teacher quality in the higher education classroom. They developed a 31-item scale to measure the level of teacher quality in the classroom. While using a factor analysis approaches, Spooren et al. were able to identify 10 factor groupings in which to measure the level of teacher quality in the classroom. The ten factor groups consist of (1) the clarity of objectives, (2) the value of subject matter, (3) the build-up of subject matter, (4) the presentation skills, (5) the harmony organization course – learning process, (6) the contribution to understanding the subject matter, (7) the course difficulty, (8) the help of the teacher during the learning process, (9) the authenticity of the examination(s), and (10) the formative examination(s). All ten factor groups warranted Chronbach alpha levels ranging between .73 and .88.

Figure 1. Macro-level Approach to Measuring Student Satisfaction in an Adult Learning and Development Master's Degree Program.



While these ten factors were considered during the development of the *ALDSS Scale*, the overall focus of our research was slightly different. In developing the *ALDSS Scale*, the researchers were interested in developing a survey that would measure the students' overall experience with the ALD master's degree program. The *ALDSS Scale* was designed to measure the level of student satisfaction with respect to six factor groups (see Figure 1). The six factor groups include: (1) curriculum, (2) learning format, (3) course materials, (4) program access, (5) faculty and instruction, and (6) faculty advising.

The curriculum, learning format, and course materials factor groups were designed to measure the students' satisfaction with respect to theory and professional practice. The program access factor was designed to measure the students' satisfaction with respect to their ability to participate in the adult learning and development program in relation to their work and family needs. The faculty and instruction factor was designed to measure the satisfaction the students had with the delivery of information in the classroom with respect to their adult learning needs. The faculty advising factor was developed to measure the students' satisfaction with respect to their academic advising sessions with the ALD faculty.

4. Research Methodology

The researchers will discuss the process by which this study was developed and administered in order to measure the level of reliability between the six factor groups pertaining to student satisfaction. This section will discuss three areas of research methodology, which include: the instrumentation, the data analysis, and the sample. The researchers also will discuss these three areas of research methodology with respect to theory and practice.

4.1 Instrumentation

The researchers used brainstorming techniques to develop the initial survey. The initial item pool consisted of 60-items. In crafting the survey items and rate scale, the researchers followed the instrument development methods described by Converse and Presser (1986) and Spector (1992). The survey was pre-piloted and reviewed by four faculty members in the area of adult education.

It was determined through the pre-pilot process that the survey consisted of items that could be categorized into six distinct groups. In addition to sorting the items into distinct groups, the pre-pilot session allowed the researchers to delete the redundant items. The researchers made a consistent attempt to develop a survey that warranted an equal number of items for each distinct group. Therefore, the researchers were able to settle on a 6-item scale for each of the six groups.

The *ALDSS Scale* consists of a 36-item scale, shown later in Table 1. The items for the *ALDSS Scale* are grouped on the survey with respect to the following six factor groups: (1) curriculum, (2) learning format, (3) course materials, (4) program access, (5) faculty and instruction, and (6) faculty advising. In addition, all of the items on the *ALDSS Scale* are equally distributed among the six factor groups. For example, items 1-6 represent the curriculum factor, items 7-12 represent the learning format factor, items 13-18 represent the course materials factor, items 19-24 represent the program access factor, items 25-30 represent the faculty and instruction factor, and items 31-36 represent the faculty advising factor.

The students were asked to rate each of the items on the *ALDSS Scale* using a 6-point Likert scale (Likert, 1932; Likert, Roslow, & Murphy, 1934) ranging from 1 = Strongly Disagree and 6 = Strongly Agree. The researchers chose to utilize a Likert scale for measuring student satisfaction because of its historical acceptance among researchers who study attitudinal data among adult learners (e.g., Ferguson, 1941; Fraenkel & Wallen, 2003; Guy & Norvell; 1977; Mueller, 1986; Spector, 1992; Spooen et al., 2007; Thurstone, 1928). One advantage in using a Likert scale is that the items representing each of the factors can be converted into a scale mean statistic. Therefore, each of the scale mean scores for each of the factor groups can be compared statistically with respect to its measureable difference. According to Spooen et al. (2007), a second “advantage of Likert scales is the ability to test each scale for reliability by means of the Cronbach’s alpha statistic” (p. 670). The 6-point rating scale was chosen by the researchers, rather than the normal 5-point scale, because it requires the participants to rate either a positive or negative response for each of the scale items.

4.2 Data Analysis

The item validation process occurred throughout the development of the *ALDSS Scale*. On two different occasions, the researchers piloted the *ALDSS Scale* to five ALD students in order to test for the content validity (Borg, Gall, & Gall, 1993) and construct validity (Carmines & Zeller, 1979) of the items with respect to the six factor groups. Each of the 36-items was typed on a single strip of paper. The students were given all 36 strips of paper and asked to read and place each item-strip into one of six envelopes representing the six factor groups. According to Carmines and Zeller (1979, p. 27), “construct validation focuses on the extent to which a measure performs in accordance with theoretical expectations.” Therefore, when items were placed into the wrong envelope, the researchers consulted with the validation participants to better understand the way in which they interpreted the item in question. The researchers either reworded the item or eliminated the item from the *ALDSS Scale*. When an item was deleted for a particular factor group, the researchers developed a new item for the factor group. The new item was based upon the recommendations provided by the ALD students participating in the validation study. The validation process took two months to complete.

The data from the students who participated in the *ALDSS Scale* was entered into an SPSS 14.0 (SPSS Inc., 2005) dataset. The rate of reliability among the six factor groups was measured using the internal consistency method. The internal consistency method was employed using the Cronbach alpha coefficient (Cronbach, 1951). According to Fraenkel and Wallen (2003) and Spector (1992), an acceptable Cronbach alpha coefficient for most studies administered in educational settings is a reliability coefficient of .70 or higher. However, due to the moderate sample size for this study, the researchers elected to use a Cronbach alpha coefficient of .80 or higher as an acceptable rate of reliability.

While testing for the rate of reliability among the six factor groups, the researchers looked at the Cronbach alpha coefficient for each factor when specific items are deleted from the study. When measuring the adjusted Cronbach alpha coefficient for factors with deleted items, the SPSS software evaluates the inter-item correlation within each of the six factor groups using a Pearson’s Correlation statistic to make the adjustment for the Cronbach alpha coefficient. The researchers also used the Pearson’s Correlation statistic to measure the inter correlation between the six factor groups.

Finally, the researchers used the Cronbach alpha coefficient to measure the reliability among the six factor groups as it pertained to race. The researchers ran independent measurements for the White students and for the Students-of-Color. The purpose for measuring the reliability coefficients with respect to race was to be able to determine if the *ALDSS Scale* was free of any scale item bias (e.g., Huebner & Dew, 1993; Plant & Devine, 1998; Tetlock & Mitchell, 2008). The researchers also would have desired to study the rate of reliability with respect to gender, but the extreme differences in sample size between the male and female students prevented them from looking at gender bias issues, as women were the majority population in the sample.

4.3 Sample

The sample size for this study consisted of 116 ALD graduate students nearing graduation, which warranted a 93.5% rate of response. It was believed that these students would provide a more reliable critique of the six factors groups regarding their satisfaction with the ALD program than a student who was new to the academic program. Whereas, a new student would not be able to rate some of the items on the *ALDSS Scale*. All student participation was voluntary and their identities were held with strict confidentiality. The sample for this study consists of a very diverse group of adult learners. The ALD students had a mean age of 38.4 years, who ranged between 23-59 years. The sample also was a very experienced group of adult learners, whose level of professional experience had a mean score of 11.0 years and a ranged between 0-36 years of experience. This study consists of a gender skewed sample, with 89.7% of the participants (n=104) representing female adult learners. In addition, the sample was racially diverse in that 53.4% of the adult learners represented students-of-color, with 50.0% African-American (n=58) and 3.4% Hispanic/Latino (n=4). The majority of the sample (62.9%) was non-married adult learners (n=73), which consisted of those who were single (n=61), divorced (n=11), and widowed (n=1). In addition, the majority of this sample (51.7%) did not have any dependent children. Although, the adult learners who had dependent children ranged between 1-4 children, with the majority consisting of adult learners with 1-child (n=23) and 2-children (n=26). Finally, fifty percent of the adult learners (n=58) for this study had an annual salary that ranged between \$30,000-\$49,999, while 28.4% earned less than \$30,000 annually and 21.6% earned \$50,000 or more annually. The sample also consisted of 11 adult learners (9.5%) who had an annual salary below that of the national poverty level.

5. Research Findings

In this section of the paper, the researchers will discuss the results from testing the reliability measures for the six factor groups and the strength of the individual items among the six factor groups. First, the researchers will discuss the findings with respect to the first research question which measures the rate of reliability among the six factor groups. The researchers will address the second research question that measures the degree to which each item among the six factor groups either strengthens or weakens the reliability measurement for the respective factors. Finally, the researchers will discuss the findings of the third research question which measures the rate of reliability of the six factor groups with respect to the race of the sample.

5.1 Does each of the six factor groups warrant an acceptable rate of reliability?

The researchers measured the rate of reliability among the six factor groups using the Cronbach (1951) alpha statistic. The reliability measurements for each of the six factor groups were generated from a sample size of 116 students (see Table 1). As suggested earlier, the researchers used the Cronbach alpha level of .80 or greater to serve as the benchmark for determining the acceptable rates of reliability among the six factor groups.

The findings suggest that each of the six factor groups had Cronbach alpha levels greater than .80. The *learning format* factor group warranted the highest rate of reliability, with an alpha level of .91 and a factor mean of 4.89. The second highest rate of reliability measurement was found among the *faculty advising* factor group, with an alpha value of .88 and a factor mean of 5.00. The *course materials* factor group had the third highest rate of reliability statistic among the six factor groups, with an alpha value of .85 and a factor mean of 5.01. The *program access* and *curriculum* factor groups warranted alpha values of .84 and .82 and factor means of 5.12 and 5.13 respectively. The factor with the lowest reliability rating, but still acceptable, was the *faculty and instruction* factor group which warranted an alpha value of .81 and a factor mean of 5.31. One of the interesting statistics that the researchers found was that while the factor means decreased, the rates of reliability increased between each of the six factor groups.

5.2 Does the rate of reliability increase among any of the six factor groups when specific items are removed?

The purpose of this question was to test whether item deletion on the *ALDSS Scale* resulted in an increase in the rate of reliability among its respective factor group. As in the case of the first research question, this research question was tested using a sample size of 116 students. The findings suggest that a decrease in the rate of reliability would occur if any items were deleted among the *curriculum* ($\alpha = .78$ to $.81$), *learning format* ($\alpha = .89$ to $.90$), and *program access* ($\alpha = .80$ to $.82$) factor groups. The findings suggest that the rate of reliability for the *faculty advising* factor group would either remain the same or decrease, if any of the six items were deleted. Deleting any of the six items would result in a Cronbach alpha range between $.84$ and $.88$. There were two items, if deleted from the *ALDSS Scale*, would have resulted in the *faculty advising* factor group hold the current Cronbach alpha value of $.88$. The first item in question was Item-31, which read “*I was able to schedule academic appointments with my advisor in a timely manner.*” The second item in question was Item-34, which read “*My advisor suggested I take specific elective courses that would prepare me for professional practice.*”

The findings suggest one item among the *course materials* factor group and one item among the *faculty and instruction* factor group are called into question. With regards to the *course materials* factor group, if either one of Items 13-17 are deleted, then this would result in a decrease in the rate of reliability ($\alpha = .80$ to $.82$) from the factor group’s current Cronbach alpha value of $.85$. However, if Item-18 (“*The course syllabi were clear with respect to the course objectives.*”) was deleted from the *course materials* factor group, then this would increase the rate of reliability to a Cronbach alpha value of $.88$. It is also important to note that the findings suggest that Item-18 was the highest rated item among the *course materials* factor group, with an item mean score of 5.33 . In looking at the inter-item correlation data, the findings suggest that Item-18 warranted a Pearson’s correlation statistic greater than a 95 percent level of confidence with respect to the five remaining items in the *course materials* factor group.

With respect to the *faculty and instruction* factor group, if either one of Items 25-29 are deleted, then this would result in a decrease in the rate of reliability ($\alpha = .74$ to $.78$) from the factor group’s current Cronbach alpha value of $.81$. In contrast, if Item-30 (“*The ALD faculty would participate frequently in the web-based course discussions for on-line course.*”) was deleted from the *faculty and instruction* factor group, then this would increase the rate of reliability to a Cronbach alpha value of $.87$. It is important to mention that the findings suggest that Item-30 was the lowest rated item among the *faculty and instruction* factor group, with an item mean score of 4.71 . The findings from the inter-item correlation data suggest that Item-30 warranted a Pearson’s correlation statistic greater than a 95 percent level of confidence with respect to the five remaining items in the *faculty and instruction* factor group.

5.3 Does each of the six factor groups warrant an acceptable rate of reliability with respect race?

The purpose of this research question was to test whether there was any racial bias regarding reliability measures among the six factor groups for the *ALDSS Scale* (see Table 2). This was accomplished by looking at the White student sample and the Students-of-Color sample independently with respect to construct reliability testing. According to the findings, the following four factor groups: (1) *curriculum*, (4) *course materials*, (3) *program access*, and (4) *faculty advising* had greater Cronbach alpha scores, ranging between $.84$ and $.90$, among the Students-of-Color sample than for the White student sample who had Cronbach alpha scores that ranged between $.79$ and $.87$ for the same four factor groups. In the case of the *learning format* factor group, the findings suggest that there was no difference in the reliability scores between the two sample groups, as both racial groups warranted a Cronbach alpha score of $.91$, which was the same value for the total sample. In contrast to the other factor groups, the findings suggest that the rate of reliability for the *faculty and instruction* factor group was much greater among the White student sample ($\alpha = .85$) than for the Students-of-Color sample ($\alpha = .68$). In addition to the *faculty and instruction* factor group not warranting an acceptable Chronbach alpha score of $.70$ among the Students-of-Color sample, the findings also suggest that there is likely a strong racial bias among the items measuring *faculty and instruction* in favor of the White student sample.

When the researchers looked further at the internal consistency of the six items measuring *faculty and instruction*, the findings suggests again that if Item-30 (“*The ALD faculty would participate frequently in the web-based course discussions for on-line course.*”) was deleted from the *ALDSS Scale* that the rate of reliability for this factor group would increase for both sample groups.

The findings suggested that if Item-30 was deleted that the *faculty and instruction* factor group would warrant a Cronbach alpha score of .77 among the Students-of-Color sample and a Cronbach alpha score of .90 among the White student sample. Although, deleting Item-30 would increase the rate of reliability among both samples, the difference in the Cronbach alpha scores for the *faculty and instruction* factor group only reduced from .17 to .13 points between the two samples. The findings also suggest that if any of the five remaining items (Items 25-29) measuring *faculty and instruction* were deleted; this would result in a decrease in the Cronbach alpha score for this factor group among both samples.

Table 1. The Measurement of Reliability for the Six Factor Groups and the Influence of Item Deletion (N=116).

I. Curriculum: Factor Mean = 5.13 SD = 0.81 Cronbach Alpha = .82		
Items	Current Item Mean	Cronbach Alpha if Item Deleted
1-The core ALD courses (ALD 605, 607, 663, 645, and 688) increased my knowledge of the adult learning and development theories.	5.47	.81
2-The core ALD courses were applicable to my professional practice.	5.07	.78
3-The elective courses were applicable to my professional practice.	4.99	.78
4-The elective courses represented my areas of interests in professional practice.	4.96	.81
5-The ALD internship course increased my skills for professional practice.	5.23	.78
6-The ALD internship course increased my theoretical framework with respect to my professional practice.	5.04	.79
II. Learning Format: Factor Mean = 4.89 SD = 0.86 Cronbach Alpha = .91		
Items	Current Item Mean	Cronbach Alpha if Item Deleted
7-The course lectures were applicable to my professional practice.	4.84	.89
8-The in-class group discussions were applicable to my professional practice.	5.02	.90
9-The group projects were applicable to my professional practice.	4.84	.89
10-The course papers were applicable to my professional practice.	5.01	.89
11-The student presentations were applicable to my professional practice.	4.78	.90
12-The course media (e.g., films, videos, and other productions) were applicable to my professional practice.	4.85	.90
III. Course Materials: Factor Mean = 5.01 SD = 0.72 Cronbach Alpha = .85		
Items	Current Item Mean	Cronbach Alpha if Item Deleted
13-The textbooks increased my knowledge of the adult learning and development theories.	5.04	.82
14-The textbooks were applicable to my professional practice.	4.79	.81
15-The additional course readings increased my knowledge of the adult learning and development theories.	5.14	.81
16-The additional course readings were applicable to my professional practice.	4.84	.80
17-The course handouts were applicable to my professional practice.	4.91	.82
18-The course syllabi were clear with respect to the course objectives.	5.33	.88*
IV. Program Access: Factor Mean = 5.12 SD = 0.92 Cronbach Alpha = .84		
Items	Current Item Mean	Cronbach Alpha if Item Deleted
19-The ALD courses were offered at a time of day that did not conflict with my work schedule.	5.16	.81
20-The ALD courses were offered at a time of day that allowed me to work around my family needs.	5.19	.80
21-The ALD core courses were offered frequently enough to meet my academic requirements.	4.96	.81
22-The elective ALD courses were offered frequently enough to meet my academic needs.	4.51	.82
23-The web-based ALD courses allowed me to participate in coursework that did not conflict with my work schedule.	5.51	.81
24-The web-based ALD courses allowed me to participate in coursework that did not conflict with my family needs.	5.40	.81

*Denotes an increase in the Cronbach alpha value for a factor if a specific item is deleted.

Table 1. The Measurement of Reliability for the Six Factor Groups and the Influence of Item Deletion (N=116). (Continued)

V. Faculty and Instruction: Factor Mean = 5.31 SD = 0.71 Cronbach Alpha = .81		
Items	Current Item Mean	Cronbach Alpha if Item Deleted
25-The ALD faculty had a strong understanding of the course topics.	5.62	.77
26-The ALD faculty stimulated my interests in the field of adult learning and development.	5.44	.75
27-The ALD faculty clearly stated the requirements for each course.	5.52	.78
28-The ALD faculty offered multiple modes of learning that met my adult learning needs.	5.32	.76
29-The ALD faculty provided a quality critique of my coursework.	5.25	.74
30-The ALD faculty would participate frequently in the web-based course discussions for on-line courses.	4.71	.87*
VI. Faculty Advising: Factor Mean = 5.00 SD = 1.06 Cronbach Alpha = .88		
Items	Current Item Mean	Cronbach Alpha if Item Deleted
31-I was able to schedule academic appointments with my advisor in a timely manner.	5.18	.88
32-My advisor responded to telephone/email messages in a timely manner.	5.16	.87
33-My advisor effectively explained the requirements for the ALD master's degree program.	5.18	.84
34-My advisor suggested I take specific elective courses that would prepare me for professional practice.	4.55	.88
35-My advisor effectively explained the requirements for the ALD internship course.	4.78	.86
36-My advisor effectively explained the ALD exit strategy requirement.	5.16	.85

*Denotes an increase in the Cronbach alpha value for a factor if a specific item is deleted.

Table 2. The Measurement of Reliability for the Six Factor Groups with Respect to Race (N=116).

Factor Groups		Total Sample Cronbach Alpha (N=116)	White Students Cronbach Alpha (n=54)	Students-of-Color Cronbach Alpha (n=62)
I.	Curriculum	.81	.79	.84
II.	Learning Format	.91	.91	.91
III.	Course Materials	.85	.84	.86
IV.	Program Access	.84	.80	.86
V.	Faculty and Instruction	.81	.85	.68
VI.	Faculty Advising	.88	.87	.90

6. Discussion

In reflecting upon the findings of this study, the researchers believe that there are a number of positive outcomes due to the development of the *ALDSS Scale*. One of the most obvious aspects to occur is that this study suggests that all six factor groups (*curriculum, learning format, course materials, program access, faculty and instruction, & faculty advising*) for the *ALDSS Scale* were found to be reliable factors. Secondly, the most important outcome stemming from the development of the *ALDSS Scale* is that the adult learning and development faculty have a reliable mechanism for measuring how satisfied their students are with the ALD master's degree program. The *ALDSS Scale* has become a very powerful tool in that the ALD faculty now has a working model (see Figure 1) for measuring ALD program's level of success with respect to numerous macro-level factors. Often when we think of program evaluation in higher education, we tend to look independently at the numerous micro-level factors, such as: teaching methods, adult learning gains, grade-point averages, graduation rates, graduate employment rates, and total student credit hours generated (e.g., enrollment rates and tuition revenue). While these are valid factors to study, most of the factors are not directly influenced by the university faculty. In contrast, the researchers developed the *ALDSS Scale* to measure six factors that summarized the overall administration of the ALD master's degree program and in which the faculty had a great deal of control over the total administration of these factors.

In essence, the *ALDSS Scale* is as much of a measurement tool for reflecting faculty performance, as it is for program evaluation. While the *ALDSS Scale* is a reliable instrument, this is not to say that the *ALDSS Scale* is without its limitations. For example, the findings suggested that two of the six factor groups had one item (Item-18 and Item-30) that reduced the potential rate of reliability for the factor groups by .03 and .06 points respectively. The *course materials* factor and the *faculty and instruction* factor did have reliability rates above .80, but anytime specific items are shown to reduce the rate of reliability by .03 points or more, one must take a second look at these items.

With regard to Item-18 for the *course materials* factor (“*The course syllabi were clear with respect to the course objectives.*”), the researchers believe that this is a valid item with respect to its application to the factor group. The researchers’ belief was reassured from the findings from the Pearson’s correlation, which suggested that there was a statistically significant correlation between Item-18 and the other items that represent the *course materials* factor group. In addition, Item-18 was the highest rated item among the *course materials* factor. The problem regarding Item-18 is not with its validity, but with its language. The researchers believe that the words “clear” and “objectives” leave the students with a vague understanding of Item-18. Therefore, the researchers recommend changing Item-18 to read, “*The course syllabi identified the course requirements.*”

With regard to Item-30 for the *faculty and instruction* factor (“*The ALD faculty would participate frequently in the web-based course discussions for on-line courses.*”), the researchers also believe that this is a valid item with respect to its application to the factor group. As in the case for Item-18, the findings from the Pearson’s correlation suggested that there was a statistically significant inter-item correlation between Item-30 and the other items that represent the *faculty and instruction* factor group. Therefore, to immediately delete this item would pose a problem of appearing to skew the results, because Item-30 was by far the lowest rated item among the *faculty and instruction* factor group. The researchers on the surface believed that the phrase “would participate frequently” causes Item-30 to read less clear. Therefore, the researchers are initially recommending that Item-30 be reworded to read, “*The ALD faculty participated in the course discussions for on-line courses.*” However, when the researchers sought to test the third research question that looked at whether there was any racial bias among the items for the six factor groups, a more serious problem developed with respect to Item-30.

When the researchers retested the Cronbach alpha levels among the six factor groups independently with respect to race, the findings suggested that five of the factor groups warranted greater Cronbach alpha levels among the Students-of-Color sample than for the White student sample. This was a very important finding, because it illustrated that there was no bias with respect to race among the items that measure the following five factor groups: *curriculum, learning format, course materials, program access, and faculty advising*. Given that the majority of the ALD graduate students are Students-of-Color, this finding suggest that when we were developing the items for the *ALDSS Scale* that we were at least cognizant of the varying cultural needs and interests with respect to our adult learners.

While these positive findings regarding construct racial bias were exciting news, we cannot ignore the negative findings that suggest that there was a strong racial bias with respect to the *faculty and instruction* factor group items in favor of the White student sample (e.g., Huebner & Dew, 1993; Plant & Devine, 1998; Tetlock & Mitchell, 2008). The findings suggested that if Item-30 were deleted, then the *faculty and instruction* factor group would warrant an acceptable Cronbach alpha level among the Students-of-Color sample as well as further increase the already acceptable Cronbach alpha level for the White student sample. Therefore, given the racial bias issue with regards to this item, the researchers recommend that Item-30 be definitely dropped from the *ALDSS Scale*, regardless of any previous concerns that the researchers may have had regarding the deletion of this item. Nevertheless, a silent issue still hovers over the *faculty and instruction* factor group. While we can solve the unacceptable Cronbach alpha level for the faculty and instruction factor group among the Students-of-Color sample by deleting Item-30, we have only decreased the disparity between the two sample groups’ Chronbach alpha levels by a total of .03 points. Therefore, there is still a .13 Cronbach alpha point difference between the reliability scores between the two racial groups, in favor of the White student sample.

When looking at the five remaining items (Items 25-29) measuring *faculty and instruction* issues, it is not visually apparent to the researchers as to what is influencing the racial bias with respect to the language of the five items. However, the researchers acknowledge that they are evaluating the items from our White lenses, as both researchers are not Faculty-of-Color.

Because the *faculty and instruction* factor group shows an acceptable rate of reliability (following the deletion of Item-30) for both sample groups, the researchers recommend that the five items (Items 25-29) measuring the *faculty and instruction* factor group remain within the *ALDSS Scale*. However, the researchers plan to conduct further statistical measurements to depict additional psychometric flaws with respect to the racial bias of these five items. Furthermore, the greatest identification of racial bias among the five items may only occur from an extensive qualitative investigation with ALD Students-of-Color. It is through their lenses that the missing pieces of the puzzle can be found.

According to Sandmann (2010), it is becoming more apparent that the faculty members who teach in adult education programs at the college or university level identify ways in which they can justify their academic programs' accomplishments and worth to the community. Therefore, the researchers believe that faculty and administrators in adult education graduate programs can easily transfer many of the same factors and items from the *ALDSS Scale* to their own academic program with respect to measuring student satisfaction. In fact, the inter-correlation data suggest that the six factor groups, which make up the *ALDSS Scale*, are statistically significant factors for measuring student satisfaction.

Furthermore, this investigation illustrates the need for survey developers to be cognizant of specific biases when developing psychometric instruments. Too often when we develop questionnaires for quantitative research, we tend to look only at the surface with respect to the reliability of the measured constructs. Had the researchers not delved further into the data, they would not have discovered the racial bias among one of the factor groups on the *ALDSS Scale*. The researchers only regret not having a more inclusive sample so that they could have studied the potential gender bias among the six factor groups.

The researchers' goal is that the *ALDSS Scale* will provide adult and higher education practitioners with another means for evaluating the success of their academic programs. With the current era of higher education bound by program accountability, adult education graduate programs must be prepared to justify their existence to the decision makers within higher education. In successfully providing valid program evaluation ideas, perhaps the closing of adult education graduate programs by universities around the country can be stymied or averted. As Sandmann (2010, p. 228) suggests, the call "for accountability and transparency will continue to shape higher education." Milton, Watkins, Studdard and Burch (2003) urges professional adult educators to pay attention to the larger context in which their programs exist, including the political climate, budget and resource constraints, and competing programs. As adult educators and program planners, it is important to be attuned the changes that are shaping higher education and thus graduate adult education programs.

Finally, the researchers encourage all faculty members to understand the importance of program evaluation and conduct further empirical research to further program evaluation efforts in adult education graduate programs. This is one area of study which is underserved yet vital to the continuing successful existence of adult education graduate programs within the higher education arena. Future research and discussion should further examine how to strengthen adult education graduate programs in the turbulent political context in which they exist.

7. References

- Borg, W. R., Gall, J. P., & Gall, M. D. (1993). *Applying educational research: A practical guide* (3rd ed.). White Plains, NY: Longman.
- Brand, M. (1997). Some major trends in higher education. *An address delivered by the President of Indiana University to the Economic Club of Indianapolis, Indiana* on February 7.
- Brand, M. (2000). Changing faculty roles in research institutions using pathways strategy. *Change*, 32(6), 42-45.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Sage University Paper series on the Quantitative Applications in the Social Sciences, series no. 17. Thousand Oaks, CA: Sage.
- Converse, J. M., & Presser, S. (1986). *Survey questions: Handcrafting the standardized questionnaire*. Sage University Paper series on the Quantitative Applications in the Social Sciences, series no. 63. Thousand Oaks, CA: Sage.
- CPAE (2010). *Standards for graduate programs in adult education*. Bowie, MD: Commission of Professors of Adult Education: American Association for Adult and Continuing Education.
<http://web.memberclicks.com/mc/page.do?sitePageId=85438&orgId=cpae>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334.

- Ferguson, L. W. (1941). A study of the Likert technique of attitude scale construction. *Journal of Social Psychology, 13*, 51-57.
- Fraenkel, J. R., & Wallen, N. E. (2003). *How to design and evaluate research in education* (5th ed.). New York: McGraw-Hill.
- Guy, R. F., & Norvell, M. (1977). The neutral point on a Likert scale. *Journal of Psychology, 95*, 199-204.
- Huebner, E. S., & Dew, T. (1993). An evaluation of racial bias in a life satisfaction scale. *Psychology in the Schools, 30*, 305-309.
- Knox, A. B. & Fleming, J. E. (2010). Professionalization of the field of adult and continuing education (pp. 125-134). In C.E. Kasworm, A.D. Rose, & J. M. Ross-Gordon (Eds.) *Handbook of Adult & Continuing Education (2010 Edition)*. Thousand Oaks: SAGE.
- Likert, R. (1932). A technique for measurement of attitudes. *Archives of Psychology, 22*(140), 55.
- Likert, R., Roslow, S., & Murphy, G. (1934). A simple and reliable method of scoring the Thurstone attitude scales. *Journal of Social Psychology, 5*, 228-238.
- Lounsbury, J. W., Saudargas, R. A., Gibson, L. W., & Leong, F. T. (2005). An investigation of broad and narrow personality traits in relation to general and domain-specific life satisfaction of college students. *Research in Higher Education, 46*, 707-729.
- Merriam, S. & Brockett, R. (2007). *The profession and practice of adult education. An introduction*. San Francisco: Jossey-Bass.
- Mueller, D. J. (1986). *Measuring social attitudes: A handbook for researchers and practitioners*. New York: Teachers College, Columbia University.
- Milton, J., Watkins, K. E., Studdard, S. S., & Burch, M. (2003). The ever widening gyre: Factors affecting change in adult education graduate programs in the United States. *Adult Education Quarterly 54* (1), 23-41.
- Nauta, M. M. (2007). Assessing college students' satisfaction with their academic majors. *Journal of Career Assessment, 15*, 446-462.
- Neal, J. E. (1995). Overview of policy and practice: Differences and similarities in developing higher education accountability. In G. H. Gaither (Ed.), *Assessing Performance in an Age of Accountability: Case Studies. New Directions for Higher Education* (pp. 5-10), No. 91. San Francisco, CA: Jossey-Bass.
- Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811-832.
- Ruppert, S. S. (1995). Roots and realities of state-level performance indicator systems. In G. H. Gaither (Ed.), *Assessing Performance in an Age of Accountability: Case Studies. New Directions for Higher Education* (pp. 11-23), No. 91. San Francisco, CA: Jossey-Bass.
- Sandmann, L. R. (2010). Adults in four-year colleges and universities: Moving from the margin to mainstream? In C. E. Kasworm, A. D. Rose, & J. M. Ross-Gordon (Eds.), *Handbook of Adult and Continuing Education* (2010 ed.) (pp. 221-230). Thousand Oaks, CA: Sage.
- SPSS Inc. (2005). *SPSS 14.0 software*. Chicago, IL: Author.
- Spector, P. E. (1992). *Summated rating scale construction: An introduction. Sage University Paper series on the Quantitative Applications in the Social Sciences*, series no. 82. Thousand Oaks, CA: Sage Publications.
- Spooren, P., Mortelmans, D., & Denekens, J. (2007). Student evaluation of teaching quality in higher education: Development of an instrument based on 10 Likert-scales. *Assessment & Evaluation in Higher Education, 32*, 667-679.
- Tetlock, P. E., & Mitchell, G. (2008). Calibrating prejudice in milliseconds. *Social Psychology Quarterly, 71*, 12-16.
- Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33*, 529-554.
- Wood, G. S., Jr. (1996). A code of ethics for all adult educators. *Adult Learning, 8*(2), 13-14.
- Zullig, K. J., Heubner, E. S., & Pun, S. M. (2009). Demographic correlates of domain-based life satisfaction reports of college students. *Journal of Happiness Studies, 10*, 229-238.