

Surveying New Testament survey: The impact of demographics and modality on an introductory New Testament class

Bart B. Bruehler¹

Abstract: Both adult and traditional students at Indiana Wesleyan University take an introductory New Testament course in conventional, compressed, and accelerated formats and through online and onsite settings. This wide variety of demographics and modalities raises the issues of if and how the various incarnations of this course facilitate the achievement of course and institutional learning outcomes. This investigation surveys the pre-test, post-test, and final paper scores of students in varying sections of the class. An interpretation of the findings concludes that the course positively impacts the learning of all types of students. However, it also suggests that traditional students who are more familiar with the Bible may be better served by what are typically deemed andragogical approaches while adult students who are less familiar with the Bible may be better served by incorporating what are often deemed pedagogical approaches.

Keywords: online, onsite, traditional, adult, accelerated, compressed, andragogy, pedagogy, general education, outcomes

Background and Purpose

Almost all colleges and universities have some type of core curriculum or general education courses that aim to develop students with well-rounded skills and perspectives. Many Christian or church-related colleges and universities require courses in Bible or theology as part of this curriculum with a variety of goals for students, including liberal learning, critical thinking, and spiritual transformation (Holland & Webster, 2012; Walvoord, 2008; Holtz, 2003). However, this presents a special challenge when a university engages a wide variety of students with various teaching modalities. Various theorists and researchers claim that adults learn differently than children/young adults (Knowles, 1998; Merriam & Caffarella, 1999), that different delivery modes call for different instructional techniques (Garrison, 2011; Delmarter, Gravett, Ulrich, Nysse, & Polaski, 2011), and that different course durations require different pedagogical tools (Wlodkowski & Ginsberg, 2010). Can a single set of course outcomes be appropriate and appropriately delivered to students from a wide demographic spectrum in so many different ways?

Indiana Wesleyan University (IWU) seeks to do just this. IWU is a large, Midwestern, Master's level university. It has approximately 3,000 traditional students on a campus located in Marion, IN and approximately 12,000 adult students enrolled in a variety of professional programs taking courses both online and onsite at several educational centers located in Indiana, Ohio, and Kentucky. IWU is one of the many faith-based institutions that has actively developed adult education (Wlodkowski, 2003, p. 6). As a faith-based liberal arts university, one of IWU's

¹ School of Service and Leadership, College of Adult and Professional Studies, Indiana Wesleyan University, 4201 S. Washington St., Marion, IN 46953, bart.bruehler@indwes.edu

institutional learning outcomes is that all undergraduate students will be able to “explain ethical, Christ-like attitudes, values, and worldview” and apply those same attitudes, values, and worldview in their careers and communities. Of course, this outcome is met through a variety of means, but all undergraduate students at IWU, whether traditional students on the liberal arts campus or adult students in professional programs, take BIL 102: New Testament Survey as one of the key instruments for meeting this ambitious goal of Christian formation.

Adult and traditional students at IWU are expected to differ in age and work experience (though there are exceptions), but other differences emerge as well. The traditional campus has recently increased its small diverse student population, but the adult programs have had a robust presence of diverse students for several years. Most students attending the traditional campus choose IWU because it is a Christian university and come with significant Christian background. The faith background of adult students varies widely from long-term active church members to those with little exposure to Christianity. Most 3 credit courses on the traditional campus are taught in conventional 15 week semesters with approximately 135 instructional hours, but a handful of courses (including BIL 102) are offered online over the summer in a compressed format with 135 instructional hours condensed into a 4-5 week period. The adult courses are in an accelerated format with approximately 70 instructional hours over a 5 or 6 week period.

Over the past few years, the author of this study has taught each of the varieties of New Testament Survey: accelerated adult online, accelerated adult onsite, conventional traditional onsite, and compressed traditional online. The recent creation of traditional online courses and a redesign of the adult courses generated reflection about the use of course outcomes, materials, and activities across the different modalities and the different students who take this course. This experience prompted questions about student performance in the course and its connection to the institution’s larger goals of formation in “ethical, Christ-like” perspectives. This generated a primary and complex research question: How do student demographics and course delivery modality interact and impact student achievement in an introductory New Testament class? This is followed by a second related question: Can a broad institutional outcome be met equivalently by one class taught to a wide variety of students through different modalities? Thus, the purpose of this study is to examine the impact of course modality and student demographics on the effectiveness of BIL 102: New Testament Survey, or, to take a survey of the results of New Testament Survey.

Literature Review

This study intersects several dimensions of research but is largely framed within the scholarship of teaching and learning (SoTL), because it seeks to rigorously examine the practice of teaching and the results of learning with a view to improving both (Gurung & Schwartz, 2009). It joins many others in comparing the effectiveness of online and onsite (or face-to-face) courses. A recent government report offered a meta-analysis of several studies, concluding that instruction offered online is as effective as traditional classroom instruction (U.S. Department of Education, 2010, p. 39). However, questions still exist about the validity and conclusiveness of such studies (Mersotis & Phipps, 1999), about the effect of extraneous factors (Driscoll, Jicha, Hunt, Tichavsky, & Thompson, 2012), about traditional students adjusting to online courses (Galyon, 2012), and about the dynamics of teaching biblical studies online (Delmarter, Gravett, Ulrich, Nysse, & Polaski, 2011). Online courses cannot merely replicate the techniques of onsite learning, because each modality has its own strengths and weakness when contributing to student

learning outcomes (Sussman & Dutter, 2010). While a general consensus has emerged, supported by a variety of empirical studies, that online education can be as effective as traditional onsite education, much of this research has focused on courses in professional programs like business, education, technology, and nursing (U.S. Department of Education, 2010, pp. 21-26). Further research on courses in the humanities (such as biblical studies) could confirm the effectiveness of online courses and further our understanding of how to optimize both onsite and online modalities.

Next, this study also builds upon a long line of research on the differences between educating youth and educating adults, or, the differences between pedagogy and andragogy. More attention has been paid in educational research to the teaching of traditional undergraduates than to teaching adults (Jarvis, 2004; and see the extensive discussion of traditional undergraduate development in Holland & Webster, 2012). The classic resources on andragogy come from the work of Knowles (Knowles, 1970; Knowles, 1998; Knowles, Holton, & Swanson, 1998). Knowles and others (Merriam & Caffarella, 1999; Cranton, 1992; Mezirow, 1991) have argued for fundamental differences between learning in pedagogy and andragogy such as: adult learners who are self-directed vs. young learners who are dependent on the instructor, adult learners who bring a wealth of experience to connect to their learning vs. young learners who have limited life experience, and adult learners who seek application to real-life tasks vs. young learners who seek to acquire information on a subject matter. Despite the widespread acceptance and influence of this perspective, Knowles's work has been challenged from a variety of perspectives (Jarvis, 2004, pp. 121-25). For example, Jarvis (1993; 2004, pp. 251-52) argues that relevant student experience with the subject matter is the key factor that should differentiate andragogical from pedagogical approaches (not chronological age). Others have suggested that the pedagogy vs. andragogy dichotomy is misguided and stereotypical since either approach may be appropriate given the needs of the students whether they are younger or older (Holmes & Abington-Cooper, 2000; Davenport, 1987). The use of BIL 102: New Testament Survey for both traditional students typically aged 18-19 who are often early in their college career and adult students who are older and have a wider set of experiences means that one must assess the value of the andragogy/pedagogy divide and implement appropriate approaches thoughtfully in the various incarnations of this course. Thus, this study will also analyze the performance of traditional and adult students to try to discern what approaches will best serve the wide variety of students who take this course.

Finally, this study encounters the issue of time in intensive, compressed, and accelerated courses. For the sake of clarity, this investigation will refer to a shortened section (4-5 weeks) with the traditional number of instructional hours (135) as a *compressed* section. This is the case for the online version of BIL 102 for traditional students. All of the adult courses at IWU are *accelerated*, meaning that they both contain fewer instructional hours (approximately 70) over a shorter duration (5 or 6 weeks) because of the andragogical principle that adults have the ability to learn more quickly and effectively. Both faculty and students have mixed perceptions of compressed courses, questioning whether they allow time for reflection and rigor (Lutes & Davies, 2013, pp. 21-22). Workload and compressed time for assessment tasks are primary concerns in accelerated courses (Lee & Horsfall, 2010). This is despite the fact that several studies have asserted the equal effectiveness of compressed and conventional courses in achieving learning outcomes. (Vreven & McFadden, 2007; Austin & Gustafson, 2006; Swenson, 2003). Similarly, recent studies suggest indistinguishable levels of learning between adults in accelerated formats and younger students in conventional formats (Wlodkowski, 2003).

However, once again, the differing formats call for strategic shaping of the learning environment and instructional techniques (Wlodkowski & Ginsberg, 2010). Thus, this study will compare the performance of students in conventional, compressed, and accelerated versions of BIL 102 to discern the dynamics at work in each.

Participants and Method

This study collected data on various incarnations of BIL 102: New Testament Survey taught at Indiana Wesleyan University during the academic year of 2013-2014. After obtaining Institutional Review Board approval, the author of this study taught various sections of this course. The sample included 2 sections of adult students in an accelerated online format (total of 22 students), 2 sections of adult students in an accelerated onsite format (total of 31 students) 2 sections of traditional undergraduates in a compressed online format (total of 24 students), and 1 section of traditional students in a semester long onsite format (total of 29 students). This is a total of 106 students. Recall that accelerated adult courses included approximately 70 hours of work in a 5 week period. The compressed online format for traditional students included approximately 135 hours of work in a 4-5 week period, and the onsite format for traditional students included approximately 135 hours of work in a 14 week semester. A set of assessments was shared across the courses (unit quizzes, interpretation journals, and a final paper) with some assessments targeted at the specific formats following best practices (discussion boards, group presentations, reading reports, etc.). The author created all of the course materials and facilitated each section. Since research has shown that instructor effectiveness and quality has a noticeable impact on student achievement (Frick, Chadha, Watson, & Zlatkowska, 2010; Rockoff, 2004), this diminished the effect of instructor variation and allowed more focus on the role of demographics and modality. The students were informed that they would be participating anonymously in a project to improve the delivery of introductory Bible courses through an analysis of their pre-test and post-scores as well as their final paper.

Data on the performance of the various groups were gathered from three primary sources: the pre-test, the post-test, and the final course paper. To this, one may add the key measurement of the difference between the pre-test and the post-test. The pre-test/post-test was administered through the web service Qualtrics. The test opened with a set of demographic questions that identified students as adult or traditional, what modality they were employing (online or onsite), and if they felt they were familiar or unfamiliar with the New Testament. The test itself was broken down in to questions about the historical-cultural background of the New Testaments (10 questions out of 40), questions about the contents of the New Testament (18 questions out of 40), and questions about theological perspectives in the New Testament (12 questions out of 40).² The number of questions in each category roughly represents the amount of time given to that domain in the course. Also, this pre-test/post-test did not include material covered during the additional time allowed in the non-accelerated courses (e.g. some Pauline letters and most of the General Epistles in the New Testament). The following is a sample of the questions from the pre-test/post-test:

² The pre-test/post-test ended with 2 open ended questions about perspectives on salvation in the New Testament and the steps of interpretation, but many students either answered these questions very briefly or omitted them altogether. If this test is used again in the future, it may be best to place these questions at the beginning to discourage students from skipping over them.

1. What is the term used to identify the final week of Jesus' life including his teaching, crucifixion, and resurrection?
 - a. The passion
 - b. The travel narrative
 - c. The Jerusalem chronicle
 - d. The last temptation
2. Which of these pairs of letters deals extensively with issues surrounding the return of Jesus?
 - a. 1 & 2 Corinthians
 - b. 1 & 2 Thessalonians
 - c. 1 & 2 Timothy
 - d. None of the other answers
3. The culture of surrounding the NT placed the highest public importance on a person's
 - a. Self-esteem
 - b. Wealth
 - c. Honor
 - d. None of the other answers

Each section had a nearly identical final paper with the traditional students writing a slightly longer paper to accommodate the number of hours required for a traditional credit hour load. The paper addresses higher level learning outcomes such as synthesis and application. All sections of the course focused on 5 blocks of the New Testament (Mark/Matthew, Luke-Acts, John, Paul, and Revelation). In the final paper, students had to describe three of these blocks, synthesize the perspectives on salvation found in those three blocks, and then apply that synthesis to a contemporary situation or issue. Students were told to include at least 7 passages of scripture and citations of various materials from the course including lectures and textbook readings. The final papers were graded by the author according to a single 100 point rubric with the following categories: use of course resources, description of three New Testament perspectives on salvation, the synthesis of those three perspectives, application of that synthesis, and writing/formatting.

Findings

The following analysis considers 79 paired pre-tests and post-tests and 84 final papers. Out of the 106 students, 4 pre-test/post-test scores are not considered because those students received an "F" as a final grade. 15 Students did not complete the post-test, and the scores of 8 students on the pre-test/post-test were omitted because they took less than 5 minutes to complete the 40 question test. This brings the number from 106 to 79. With regard to the papers, the 4 students with an "F" were removed from the sample as well as 4 students who did not turn in the final paper. Other paper scores were removed as well: 2 students plagiarized on the final paper, 4 students did not turn in the final paper, and 8 received an "F" on the final paper (while still managing to pass). This brings the number down from 106 to 84. The following analysis breaks down the scores in various ways. First, one can divide the results into students in online sections and students in onsite sections to examine the impact of modality. Second, one can divide the results into adult students and traditional students to examine the impact of demography. Third, on the pre-test students rated themselves as either being "familiar" or "unfamiliar" with the

contents of the New Testament in order to measure how background impacted student performance. Finally, the results will be broken down according to the four groups of students/sections: adults in online sections, adults in onsite sections, traditional students in online sections, and traditional students in onsite sections. This will provide a sense of how results in the four different formats of the course compare to each other.

The Distribution of the Data

This study examines very different settings of BIL 102 across a wide range of students. The distribution of the data in some of these cases reveals interesting dynamics and challenges for evaluating the results and their implications for teaching this course best to this variety of students. When taken as a whole, the performance of all of the students distributes approximately normally in each of the primary measurements: the pre-test, the post-test, the pre-test to post-test difference, and the final paper total. However, when broken down into the various groups (adult online, adult onsite, traditional online, and traditional onsite) variations from approximately normal distribution begin to appear.

The pre-test scores ranged from 6 to 30 (out of 40) with a mean of 15.66. These results fit with the author’s expectations of student’s performance on the pre-test. The Shapiro-Wilk result is .117 for the pre-test scores, indicating approximately normal distribution as displayed in Figure 1.

The post-test score ranged from 12 to 39 (out of 40) with a mean of 22.63. The test demonstrated reasonable validity and reliability. 33 of the 40 questions were answered correctly by a majority of the students on the post-test. Of the 7 questions not answered correctly by the majority, every question received at least 16 correct responses and some as many as 42 correct responses. The Shapiro-Wilk result is .359 for the post-test scores, indicating approximately normal distribution as shown in Figure 2.

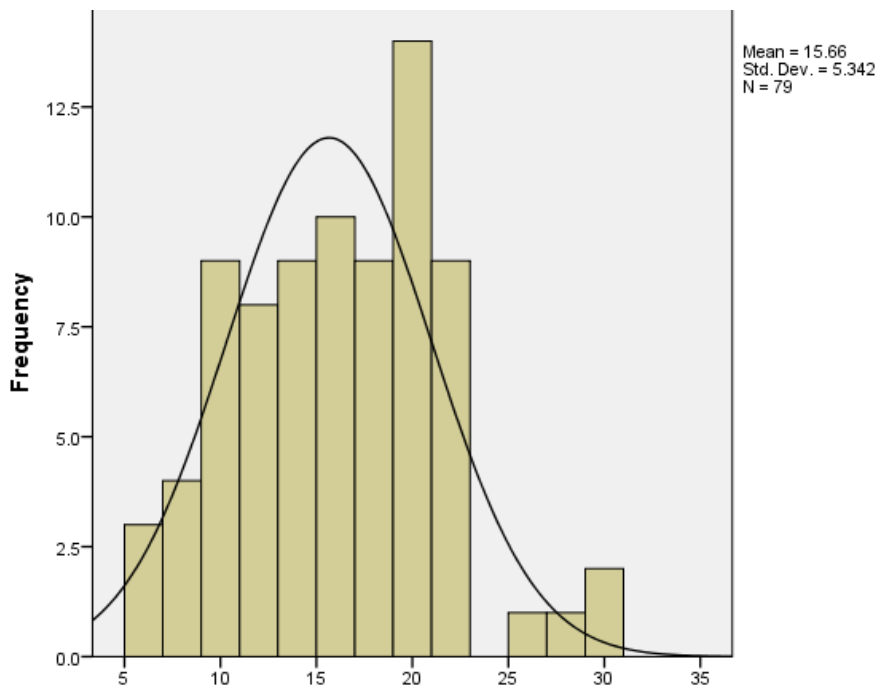


Figure 1. Pre-test scores for New Testament Survey

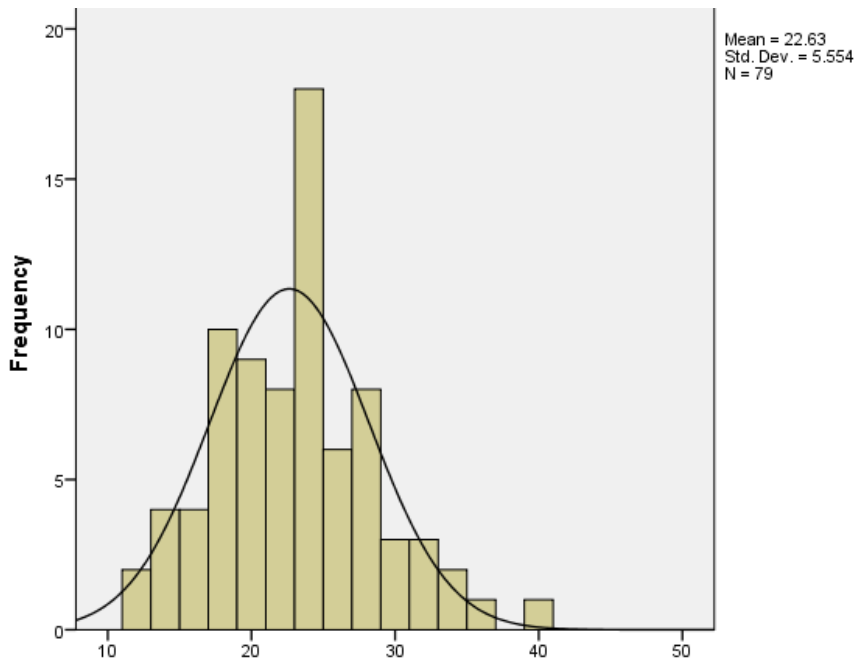


Figure 2. Post-test scores for New Testament Survey

The pre-test to post-test difference ranged from -2 to 17, which is rather large. The mean was 6.97, which is approximately 17% of the 40 available points. However, this mean also had a rather large standard deviation of 4.394. The Shapiro-Wilk test gives a result of .099 which supports an approximately normal distribution of the differences as seen in Figure 3.

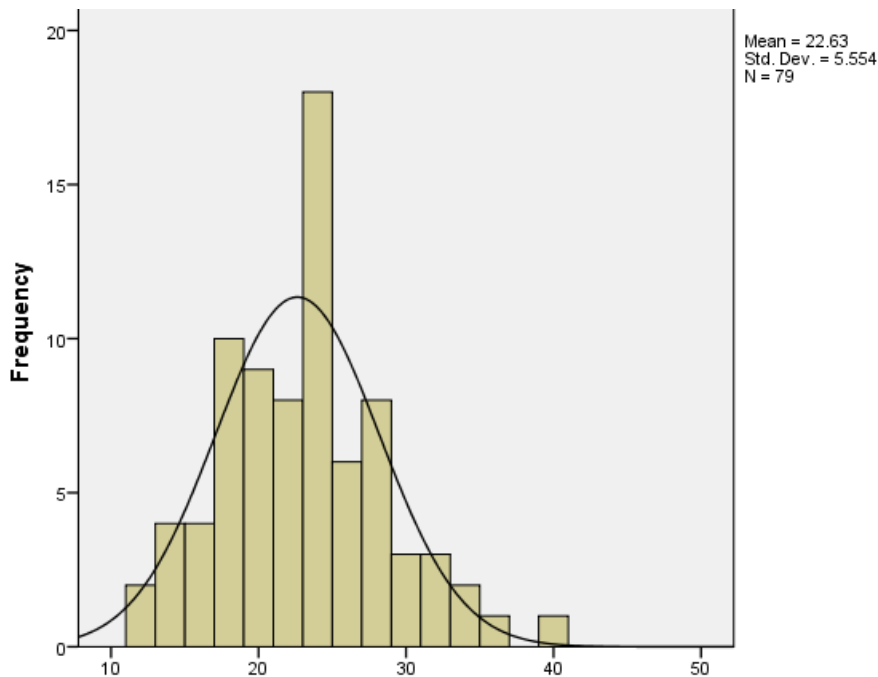


Figure 3. Pre-test to post-Test differences

The total paper score appears to be the most normally distributed of all of the data. (Note that $n = 84$ here since some students did not complete the post-test adequately). With 100 points available on the final paper, the mean is 80.71 with a standard deviation of 9.971. This means that about 68% of the scores land between 70 and 90 or that a majority of the scores ranged from a C- to a B+. Figure 4 reflects a very balanced distribution of the total paper scores.

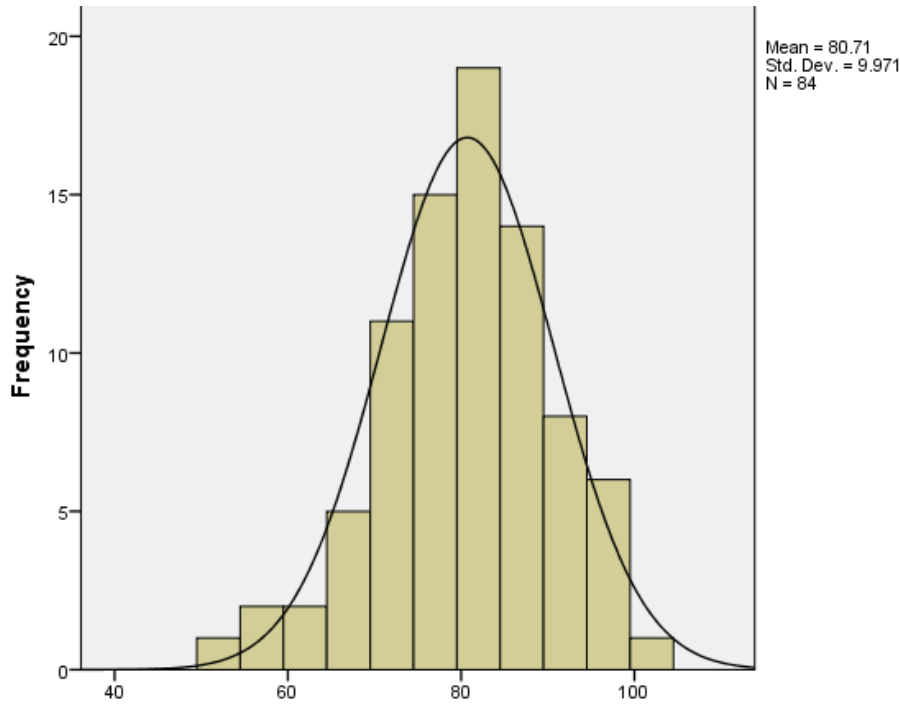


Figure 4. Total score on final paper.

However, when one breaks the data into the four separate groups (adult online, adult onsite, traditional online, and traditional onsite) some results do not distribute normally. Table 1 provides information for each of the four group’s scores on the pre-test.

Table 1

Four groups performance on the pre-test

Group	Number	Mean	Std. Dev.	Range	Shapiro-Wilk
Adult online	18	13.11	4.391	6 to 22	.438
Adult onsite	21	14.29	5.737	6 to 22	.014
Traditional online	21	16.67	4.608	9 to 27	.653
Traditional onsite	19	18.47	5.200	11 to 30	.070

Note that the Shapiro-Wilk result for the adult onsite group is .014 indicating a non-normal distribution (all other group’s scores indicate approximately normal distribution with $p > .05$). Figure 5 reveals a bi-modal distribution of the data for the adult onsite scores on the pre-test.

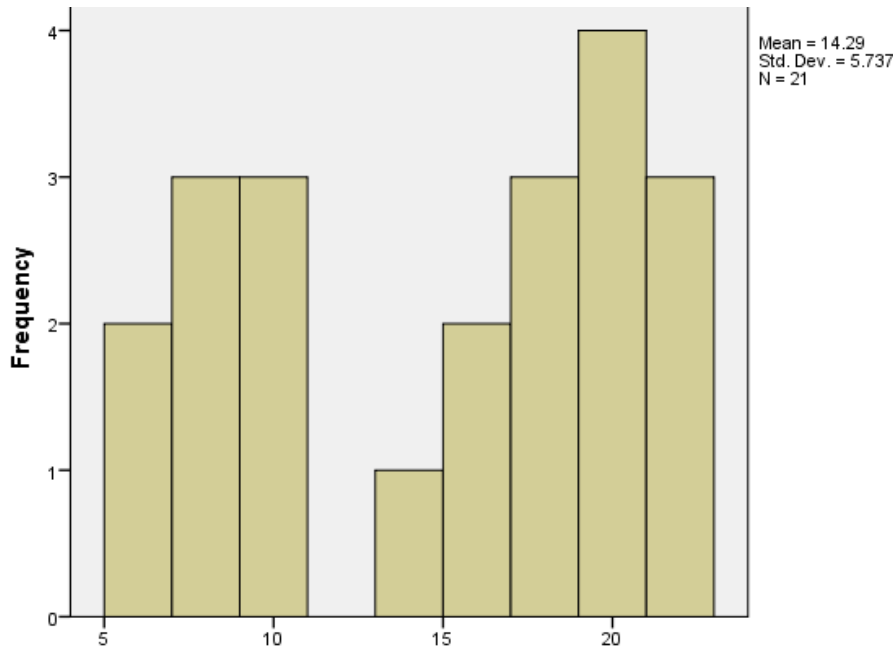


Figure 5. Adult online pre-test scores

The two peaks reflect the author’s experience teaching New Testament Survey to adult students in onsite sections. Many of these sections have two discernible groups of students: those with a long and rich Christian upbringing who are still very active in their churches and thus more familiar with the Bible and those with very little exposure to Christianity and very little familiarity with the Bible. This chart shows peaks at those very different levels as reflected in the pre-test with a few scattered in the middle. This will call for some nonparametric evaluation to compare the scores of the four groups on the pre-test.

A similar phenomenon occurs when examining the post-test scores of students in the four groups.

Table 2

Four groups post-test scores

Group	Number	Mean	Standard Dev.	Range	Shapiro-Wilk
Adult online	18	20.00	4.971	12 to 30	.663
Adult onsite	21	20.86	5.199	12 to 30	.533
Traditional online	21	24.19	4.926	18 to 34	.169
Traditional onsite	19	25.37	5.659	15 to 39	.034

Again, the scores in one of the four groups are not normally distributed, but this time it is the traditional onsite students ($p = .034$). Figure 6 reveals an extreme data point at the very top and bottom of the range with a very high peak in the lower-middle and a stretch of higher scores with no corresponding batch of lower scores.

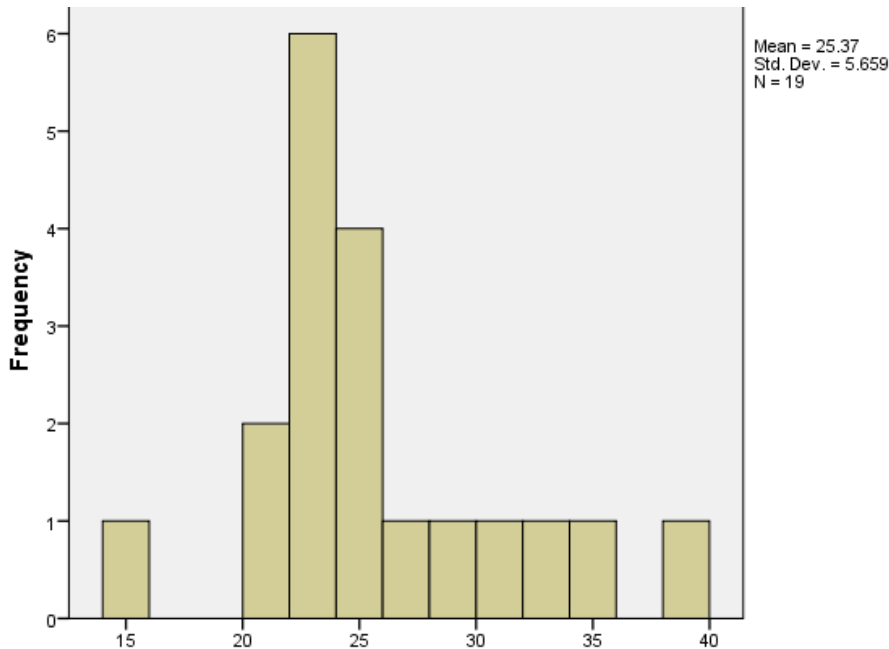


Figure 6. Traditional onsite post-test scores

This again calls for nonparametric testing when comparing the scores of the four groups on the post-test.

Table 3 provides information on the difference in scores between the pre-test and the post-test.

Table 3

Four groups pre-test to post-test difference

Groups	Mean	Standard Dev.	Range	Shapiro-Wilk
Adult online	6.89	4.337	2 to 16	.015
Adult onsite	6.57	4.823	-1 to 17	.054
Traditional online	7.52	4.523	-1 to 17	.825
Traditional onsite	6.89	4.095	-2 to 14	.833

First, it is striking that the means and standard deviations are so close across all four groups. This will be explored further below. Second, these data reveal that in three of the four groups at least one student scored worse on the post-test than on the pre-test (thus the negative score difference). Furthermore, the scores of 5 students were omitted from this data set because the web-based testing report showed that these students took less than 5 minutes to complete the 40 question post-test. Students were given credit for completing the post-test, were told not to study for it because it would not be graded, and apparently some took it with disregard, merely trying to complete the task as quickly as possible to gain the allotted points at the end of the course. Third, one of the four groups falls outside of normal distribution again as indicated by the Shapiro-Wilk score, but this time it is the adult online group ($p = .015$). Figure 7 displays an unusual distribution of differences between the pre-test and post-test among these students.

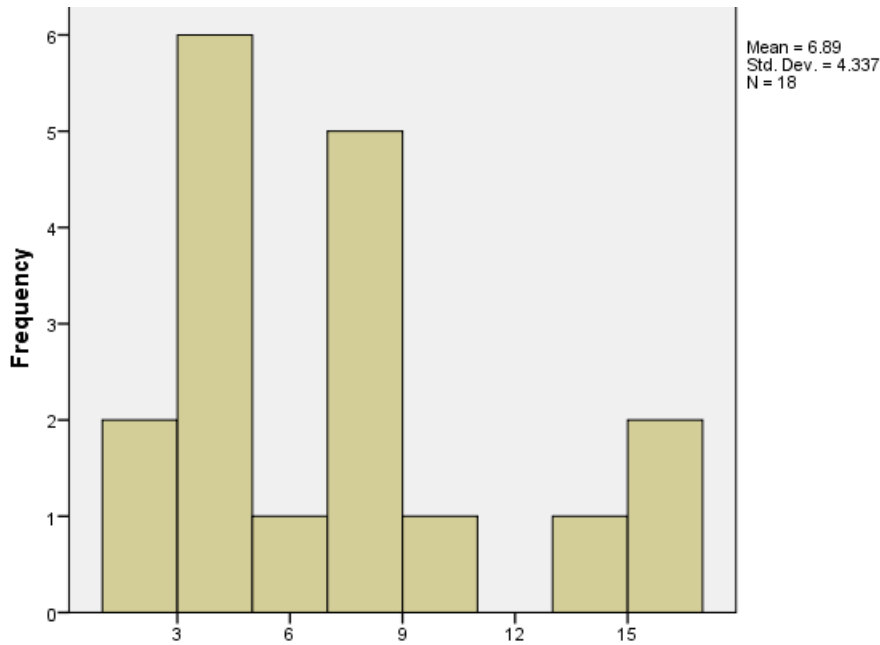


Figure 7. Adult online pre-test to post-test difference

As indicated by the students who took less than 5 minutes to complete the post-test, some of the students may have taken the post-test (or the pre-test) quickly and carelessly, seeking only to obtain the points allotted for completing it. This calls for the use of nonparametric tests when evaluating the pre-test to post-test difference among the four groups.

The data for the total paper scores are provided in Table 4.

Table 4

Four groups total paper scores

Group	Mean	Standard Dev.	Range	Shapiro-Wilk
Adult online	79.00	10.061	52 to 92	.098
Adult onsite	79.50	10.714	56 to 98	.655
Traditional online	81.50	7.990	63 to 94	.850
Traditional onsite	83.45	10.928	58 to 100	.800

The total score on the final paper is normally distributed for each group (Shapiro-Wilk score in all cases $p > .05$). One interesting feature here is that the standard deviation for the traditional online group is over 2 points lower than all the others, indicating a more compressed set of scores around the mean. The traditional online group has the highest low score and the second lowest high scores. Thus, this group had less spread and range than any of the others. Regular parametric tests can be employed when exploring the scores of the four groups on the final paper.

Pre-test Scores

The pre-test scores will first be considered according to the three sets of pairs: online and onsite students, adult and traditional students, and students who are familiar and unfamiliar with the Bible. Table 5 provides the number, mean score, and standard deviation of the scores as well as the results of the T-test for significance and the partial eta squared score for effect size.

Table 5

Paired groups pre-test scores

Group	Number	Mean	Standard Deviation	T-test	Partial Eta Squared
Online	39	15.03	4.799	.302	.014
Onsite	40	16.28	5.818		
Adult	38	13.71	5.193	.001	.125
Traditional	41	17.46	4.874		
Familiar	38	17.68	5.914	.001	.135
Unfamiliar	41	13.78	4.814		

The online and onsite student scores on the pre-test show no statistically significant difference ($p = .302$). This category cuts across both adult and traditional students focusing only on the modality of the course. The two mean scores were rather close at 15.03 and 16.28. The partial eta squared score was .014 representing a negligible effect size based on this grouping. However, separating the students into the demographic groups of adult and traditional results in strong show of statistical significance ($p = .001$). The adult student mean score was 13.71, and the traditional student mean score was 17.46, yet the effect size was only .125 which is still rather small. This may indicate that adult students tend to come into the course with a lesser degree of knowledge about the New Testament than their traditional counterparts. Other explanations are possible as well: perhaps these adult students were simply trying to complete the pre-test for the completion points, perhaps lower familiarity led to lower effort, or perhaps the adult students were poorer test takers. Very similar (and very predictable) results obtain with the familiar and unfamiliar groups. With a p value of .001 again, one can be very certain of a significant difference with the familiar student scoring almost 4 points higher on the pre-test. Students as a whole seem to have assessed their entry level knowledge of the New Testament with some degree of accuracy. However, again the effect size is on the small end of the scale at only .135. Additionally, there is a significant negative correlation between the adult-traditional and familiar-unfamiliar groupings. A Spearman's rho test shows a high degree of significance and a moderate effect size here ($r_s = -.571, p < .000$). Thus, adult students are moderately likely to be unfamiliar with the Bible (29 out of 38) while traditional students are moderately likely to be familiar with the Bible (29 out of 41). This should be remembered for the following analyses as well.

Moving on to the analysis of the four groupings, recall that the pre-test scores for the adult onsite group was not normally distributed. Thus, a Kruskal-Wallis test was run to check for significant variation across the four groups. This returned a p value of .020 supporting

statistically significant variation meriting further investigation. The post-hoc pairwise comparisons revealed that the only significant difference ($p = .024$) was between the adult online and traditional onsite groups with a difference in means of 5.36. Overall, the traditional onsite students scored higher than any of the other groups (1.80 higher than traditional online and 4.18 higher than adult onsite). This probably reflects the fact that the vast majority of traditional onsite students at the university come from Christian backgrounds and thus have greater starting familiarity with the New Testament than the population of adult students.³

Post-test Scores

The post-test scores aim to measure students' knowledge of the New Testament as they leave the course. The presentation and analysis of these scores will set the stage for analyzing the differences between the pre-test and post-test scores in the next section. As with the pre-test scores, the post-test scores will first be considered according to the three sets of pairs: online and onsite students, adult and traditional students, and students who are familiar and unfamiliar with the Bible. Table 6 provides the number, mean score, and standard deviation of the post-test scores as well as the results of the T-test for significance and the partial eta squared score for effect size.

Table 6

Paired groups post-test scores

Group	Number	Mean	Standard Deviation	T-test	Partial Eta Squared
Online	39	22.26	5.320	.555	.014
Onsite	40	23.00	5.818		
Adult	38	20.47	5.114	.001	.125
Traditional	41	24.63	5.238		
Familiar	38	24.97	5.248	.000	.135
Unfamiliar	41	20.46	4.965		

Similar to the pre-test results, the post-test shows statistically significant differences between the adult and traditional groupings ($p = .001$) and between the familiar and unfamiliar groupings ($p < .000$), but both have relatively small effect sizes ($\eta_p^2 < .2$). The traditional and the familiar students scored approximately 4 points higher than their adult and unfamiliar counterparts. The analysis of the online and onsite groupings revealed no significant difference.

When turning to the four groups, one must recall that the scores of the traditional onsite students were not distributed normally on the post-test. Thus, a Kruskal-Wallis test was run to check for significant differences, and it returned a value of .013 supporting some significant difference. Similar to the pre-test analysis, a post-hoc pairwise comparison shows a significant difference between the traditional onsite and adult online post-test scores ($p = .035$). Again, the traditional onsite students scored higher than all the other groups (5.37 more than the adult

³ The possible causes of slightly lower scores among the traditional online students are not clear from the available data.

online, 4.51 more than the adult onsite, but only 1.18 more than the traditional online). Overall, these results present a close parallel to the pre-test results: groups who scored significantly higher on the pre-test also scored significantly higher on the post-test, and the four groups stayed in the same rank order. This brings us to the next dimension of the analysis.

Pre-test to Post-test Difference

While the pre-test and post-test scores provide some interesting information about the starting and ending points of various groupings, the pre-test to post-test difference seeks to analyze the amount of learning that took place during the course. The very interesting result here is that there was no significant difference among any of the groupings' difference in scores from the pre-test to the post-test. The mean scores for all of the groupings were between 6.68 (for those unfamiliar with the Bible) and 7.52 (for the traditional online students)—less than 1 point difference overall. The p values for all of the T-tests of the paired groups (adult/traditional, online/onsite, and familiar/unfamiliar) were well over .05 and often above .80, and the effect size scores were all under .10. Recalling that the adult online differences were not normally distributed, the Kruskal-Wallis result was .735, which supports retaining the null hypothesis of no difference. As an aggregated whole all students improved from the pre-test to the post-test by approximately 7 points no matter how they might be grouped. For example the adult online students improved from approximately 13 on the pre-test to approximately 20 on the post test, and the traditional onsite students improved from approximately 18 on the pre-test to 25 on the post-test. Furthermore, no significant differences appeared when the various groups were measured by how they performed on the various sections of the pre-test/post-test (backgrounds, contents, and theology). Improvement occurred at the same rate for all groups across all the dimensions of the test.

However, the mean scores only reveal one dimension of the results. A review of the scores provided on Table 3 reveals further notable dynamics. The standard deviations and ranges in each case are rather large, especially given the mean score of approximately 7. The standard deviations for each group are larger than half the mean, pointing to the fact that 68% of the scores fell approximately between 2 and 12. Ranges also stretch rather widely from 14 points among the adult online students to 18 points for the adult onsite and traditional online students. Thus, even with a mean score of approximately 7 obtaining across all of the groupings, there appears to be a wide range of scores within each group from students with little to no improvement to those who doubled their score from the pre-test.

Paper Scores

In addition to the pre-test and post-test scores, every student wrote a nearly identical final paper. These papers were all graded on the same rubric as discussed above (resources, description, synthesis, application, and writing). Similar to the pre-test to post-test difference all significance tests with any of the groupings (e.g. online or onsite, and the four groups) returned values greater than .05, and all effect size scores were less than .10. However, some similar dynamics seen above appear here but less pronounced. The traditional students' mean score of 82.21 was 3 points higher than the adult students' mean score of 79.21 (even though $p = .169$). Also, the traditional onsite students' score of 83.45 was approximately 4 points higher than both the adult online (mean of 79.00) and adult onsite (mean of 79.50) students ($p = .510$ and $.563$

respectively). This led to a closer look at the various categories on the rubric. No significant differences were found in any of the paper categories across any of the groupings with one exception. The scores for the adult and traditional students were significantly different in the resources category ($p < .000$ and $\eta_p^2 = .295$). The adult students had a mean score of 14.79 and the traditional students had a mean score of 18.26 in this category. Thus, the statistically insignificant difference in the mean total paper scores noted above seems to be explained primarily by a statistically significant difference in how well adult and traditional students employed course resources in their papers.

Discussion and Implications

This investigation studied a single course employing different modalities and aimed at students from a range of demographic backgrounds. Thus, it is interesting that the scores for the pre-test, post-test, and paper were all distributed normally. This diverse sample of students and courses still resulted in fairly predictable distributions. However, within some of the groups diversity emerged that generated non-normal distribution. The adult onsite students' scores on the pre-test were distributed bimodally, probably indicating the presence of a few students with strong Christian backgrounds alongside others with little exposure to the Bible. The traditional onsite student scores also did not distribute normally largely due to the impact of a handful of very high performing students. This points to variation that may be due to individual aptitude or achievement. Finally, the adult online difference between the pre-test and post-test also did not distribute normally. However, this may belie a weakness in the study methodology since some online students seemed to take the post-test quickly just to earn the completion points. The paper totals were distributed normally across all the groups and a majority of final grades fell between C- and B+. The very even and balanced results of this broad and summative assessment indicate good achievement of course outcomes across all the categories.

Several conclusions follow from this analysis based on both the presence and absence of significant results. This sample confirms that adult and traditional students enter IWU with notably different levels of knowledge about the New Testament. We saw a moderately strong correlation with adult students identifying themselves as unfamiliar with the Bible and the opposite with traditional students identifying themselves as familiar with the Bible. This is borne out in a significant difference with the pre-test results as well where traditional students scored approximately 4 points higher. This raises the question about the approach and contents of New Testament Survey. The accelerated sections cover less of the New Testament than their conventional length counterparts for traditional students because of the reduced instructional time. Thus, adult students enter the course knowing less about the New Testament and leave being exposed to less of the New Testament in comparison with the traditional students. Since experience in the subject matter is a key factor in determining whether to use an andragogical or pedagogical approach (Jarvis, 1993; 2004), the traditional students' familiarity with the Bible probably merits more of an andragogical approach building on their knowledge of scripture and developing the skills of analysis and application while the adult students' lack of familiarity calls for a more pedagogical approach to increase their knowledge base and emerging interpretive skills. Implementing these could improve in the various versions of New Testament Survey and result in increased learning for both sets of students

As noted above, across all of the various categories of analysis students improved from the pre-test to the post-test by a mean score of approximately 7 points, regardless of their initial

score on the pre-test. This is a fascinatingly consistent result given the different starting points, demographic backgrounds, and modalities employed in New Testament Survey across IWU. However, it also raises further issues. First, while all students improve at nearly the same rate, final scores are definitively impacted by the student's starting point. Second, the traditional onsite students must recall instruction from over a 14 week period while all the other groupings are recalling information from only a 4-5 week period. This may point to a higher degree of sense and meaning leading to more enduring long term memory storage (Sousa, 2001, pp. 46-51). Third, this result means that adult students and those unfamiliar with the Bible (recalling that these two groupings have significant overlap) end up *leaving* the course with only a slightly higher score (about 2-3 points) than where traditional students and those familiar with the Bible *began* the course. The improvement is significant and important but it should be considered if this is adequate to help adult/unfamiliar students reach the institutional outcomes of "ethical and Christ-like values and worldview." Fourth, the standard deviation and the range of the pre-test to post-test difference were both rather large. This seems to indicate a broad array of individual achievement. Some students excelled and some students performed poorly across all the various groupings. This may indicate that student aptitude and effort may play a large role regardless of demographics or modality.

The results on the total paper score were also very similar across all the groupings with no significant differences. Thus, neither demographics nor modality had any noticeable impact as students wrote a summative paper describing, synthesizing, and applying the major theological perspectives of the New Testament covered in the course. It is interesting, however, that the adult students with their presumed emphasis on real-life application of knowledge (Knowles, 1970; 1998) performed no better than their traditional counterparts in applying New Testament material to relevant, contemporary circumstances. This is even more startling given the fact that several learning activities unique to the adult accelerated sections specifically aimed to facilitate connections between the New Testament and life or career of students as part of good andragogical practice (e.g. group presentations). Again, this may suggest that the andragogy/pedagogy distinction may not hold up so well here because of the different starting points of the traditional and adult students. Traditional/familiar students may be better prepared for the move to application while adult/unfamiliar students may need more basic instruction for comprehension. Finally, the one significant difference on the final papers was the significantly better use of course resources (Bible, textbook, and lectures) by traditional students. The approximately 65 hours of additional instructional time for the traditional students may have provided them the opportunity for broader and deeper interaction with the course materials that enabled them to use more of those materials in more relevant and appropriate ways on the final paper.

These results have relevance to other courses, especially introductory and general education courses, beyond the specific focus on the New Testament. This study suggests that prior familiarity with the content is a key factor in deciding when and how to aim at lower level learning outcomes with more "pedagogical" strategies or higher level learning outcomes with more "andragogical" strategies. Placement tests could help group students in more appropriate sections, or instructors could survey students at the beginning of the course as a guide to selecting the most beneficial materials and teaching strategies. The results raise issues about the benchmarking of broader outcomes when dealing with such a wide variety of students. Larger outcomes at the program and institutional level might do well both to establish a minimum degree of accomplishment (e.g. all students who graduate from IWU can do . . .) and to measure

the degree of growth over time (e.g. students who graduate from IWU tend to grow in these ways . . .). The minimum standard would certify the gains of students entering with lower levels of familiarity or aptitude and the documented growth would celebrate the work that the university does with students who have more background and capacity. Next, this study questions the andragogical principle that adult learners should be taught with methods stressing application. If application is addressed prematurely before more basic learning outcomes are met even in adults, then that instructional effort may be less effective and could be better used on scaffolded tasks appropriate to their level of familiarity and skill no matter what the content area. Finally, while accelerated classes have shown effective achievement of course outcomes, students unfamiliar with the subject matter will have a steeper learning curve with materials and textbooks, and this should be accounted for in class planning.

Limitations and Further Research

This investigation includes a limitation that faces many studies of classroom performance—the subjects were not randomly selected and placed into various sections of BIL 102, and this may skew the results due to issues of self-selection. However, one of the aims of this study was to see if different materials and approaches specifically targeted to a particular group of students worked equally well in achieving course outcomes across the various incarnations of the course. Another limitation emerged in the analysis of the post-tests. Some students clicked through the test in less than 5 minutes in order to earn the completion points. This may reflect a degree of carelessness on the part of some students who were seeking to check off a task, and this may have lowered the overall scores on the post-test. The pre-test/post-test was also somewhat biased toward the adult students in that it did not ask questions about material that was covered in the conventional and compressed sections but simply could not be fit into the time allotted for the accelerated sections (e.g. additional information about the Gospels, additional Pauline letters, and some General Epistles). A more comprehensive test could be used to explore the total learning of the traditional students in comparison with the adults in accelerated courses.

The study design set the author as the creator and instructor of all the courses. This offered some control on the variables of content and instructor performance, but it also brings certain limitations. The single instructor may have produced a levelling effect on scores across the courses that would not be seen in a wider sample of students with differing instructors. Future investigations could include more students, instructors, and sections to see if similar dynamics of performance across the various modalities and demographics obtain when not controlling for the instructor variable. Also, adding other raters on the final papers could provide greater objectivity and validity to these scores. Such a broader analysis would help to measure the effectiveness of this general education course across the institution.

Finally, as this study surveyed the impact of demographics and modality on the achievement of students in New Testament Survey, it raised the issue of how learning in this introductory course helps students to reach the larger institutional learning outcome of explaining, demonstrating, and applying “ethical, Christ-like values and worldview.” The pre-test to post-test improvement across all categories reflects increased knowledge about the contents of scripture, but the adult/unfamiliar students still lagged behind their traditional/familiar counterparts. This raises two questions. What is the relationship of knowledge about the New Testament and the development of and “ethical, Christ-like

worldview,” and do the lower (yet improved) scores of the adult/unfamiliar students reflect adequate progress toward this goal? The final paper components of theological synthesis and application link more directly to the goals of Christ-like values and character, and all categories of students demonstrated solid performance in these areas. Yet, adult/unfamiliar students typically appear have more ground to traverse than their traditional/familiar counterparts to reach these higher order results. This raises the following question: can IWU employ the various modalities with greater sensitivity to demographics to capitalize on the background of traditional/familiar students and intensively serve the growth of adult/unfamiliar students? New Testament Survey is the one undergraduate course at IWU that cuts across modalities and demographics to move all students further toward “ethical, Christ-like values and worldview.” Counterintuitively, employing approaches associated with andragogy for younger but more biblically literate traditional students and approaches associated with pedagogy for older but less biblically literate adult students may serve both groups better in their progress towards this goal.

Acknowledgments

The author would like to thank the Lilly Foundation for providing funds to support the Lilly Release Load Grant at IWU. This grant provided the time needed to analyze and write up the results of the study. Financial support for this study was provided by School of Service and Leadership at IWU. Finally, the author would like to thank Dr. Jim Freemyer for his helpful input on the statistical analysis.

References

- Austin, M., & Gustafson, L. (2006). Impact of course length on student learning. *Journal of Economics and Finance Education*, 5(1), 26-36. Retrieved March 20, 2014 from <http://www.economics-finance.org/jefe/econ/Gustafsonabs.pdf>.
- Cranton, P. (1992). *Working with adult learners*. Toronto: Wall and Emerson.
- Davenport, J. (1987). Is there any way out of the andragogy morass? *Lifelong Learning: An Omnibus of Practice and Research*, 2(3), 17-20.
- Delmarter, S., Gravett, S. L., Ulrich, D. W., Nysse, R. W., & Polaski, S. H. (2011). Teaching Biblical Studies Online. *Teaching Theology & Religion*, 14, 256-283. doi: [10.1111/j.1467-9647.2011.00719.x](https://doi.org/10.1111/j.1467-9647.2011.00719.x).
- Driscoll, A., Jicha, K., Hunt, A., Tichavsky, L., & Thompson, G. (2012). Performance and satisfaction in an online versus a face-to-face introductory sociology course. *Teaching Sociology*, 40, 312-21. doi: [10.1177/0092055X12446624](https://doi.org/10.1177/0092055X12446624).
- Frick, T. W., Chadha, R., Watson, C., & Zlatkovska, E. (2010). Improving course evaluations to improve instruction and complex learning in higher education. *Education Technology Research & Development*, 58, 115-36. doi: [10.1007/s11423-009-9131-z](https://doi.org/10.1007/s11423-009-9131-z).

Bruehler, B.B.

Galyon, C. K. (2012). Are undergraduate students ready for online learning? A comparison of online and face-to-face sections of a course. *Rural Special Education Quarterly*, 31(4), 25-31.

Garrison, D. R. (2011). *E-learning in the 21st century* (2nd ed.). New York: Routledge.

Gurung, R. A. R., & Schwartz, B. M. (2009). *Optimizing teaching and learning: Practicing pedagogical research*. Malden, MA: Wiley-Blackwell.

Holland, G. S., & Webster, J. S. (Eds.) (2012). *Teaching the bible in the liberal arts classroom*. Sheffield: Sheffield Phoenix Press.

Holmes, G., & Abington-Cooper, M. (2000). Andragogy vs. pedagogy: A false dichotomy? *Journal of Technology Studies*, 26(Summer-Fall). Retrieved March 3, 2014 from <http://scholar.lib.vt.edu/ejournals/JOTS/Summer-Fall-2000/holmes.html>.

Holtz, B. W. (2003). *Textual knowledge: Teaching the bible in theory and practice*. Jewish Education Series, vol. 1. No city: Jewish Theological Seminary of America.

Jarvis, P. (1993) Pedagogy, andragogy and professional education. Paper presented at the International Conference on the Training of Adult Educators, Wadham College, Oxford.

Jarvis, P. (2004). *Adult education & lifelong learning: Theory and practice* (3rd ed.). London: Routledge Falmer.

Knowles, M. S. (1970). *The modern practice of adult education*. New York: Association Press.

Knowles, M. S. (1998). *The adult learner: A neglected species* (3rd ed.). Houston: Gulf Publishing.

Knowles, M. S., Holton, E. F., & Swanson, R. A. (1998). *The adult learner: The definitive classic in adult education and human resource development* (6th ed.). Boston: Butterworth-Heinemann.

Lee, N., & Horsfall, B. (2010). Accelerated learning: A study of faculty and student experiences. *Innovative Higher Education*, 35, 191-202. doi: [10.1007/s10755-010-9141-0](https://doi.org/10.1007/s10755-010-9141-0).

Lutes, L., & Davies, R. (2013). Comparing the rigor of compressed format courses to their regular semester counterparts. *Innovative Higher Education*, 38, 19-29. doi: [10.1007/s10755-012-9226-z](https://doi.org/10.1007/s10755-012-9226-z).

Merisotis, J., & Phipps, R. (1999). What's the difference? Outcomes of distance vs. traditional classroom-based learning. *Change: The Magazine of Higher Learning*, 31(3), 12-17. Doi: [10.1080/00091389909602685](https://doi.org/10.1080/00091389909602685).

Merriam, S. B., & Caffarella, R. (1999). *Learning in adulthood: A comprehensive guide* (2nd ed.). San Francisco: Jossey-Bass.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94, 247-52. Retrieved June 8, 2014 from <http://dx.doi.org/10.1257/0002828041302244>.juyt8

Sousa, D. A. (2001). *How the Brain Learns* (2nd ed.). Thousand Oaks, CA: Corwin Press.

Sussman, S., & Dutter, L. (2010). Comparing student learning outcomes in face-to-face and online course delivery. *Online Journal of Distance Learning Administration*, 13(4). Retrieved March 20, 2014 from http://www.westga.edu/~distance/ojdla/winter134/sussman_dutter134.html.

Swenson, C. (2003). Accelerated and traditional formats: Using learning as a criterion for quality. *New Directions for Adult and Continuing Education*, 97, 83-92. Doi: [10.1002/ace.91](https://doi.org/10.1002/ace.91).

U. S. Department of Education, Office of Planning, Evaluation, and Policy Development (2010). *Evaluation of evidence-based practices in online learning: a meta-analysis and review of online learning studies*. Washington, DC: U. S. Government Printing Office. Retrieved March 10, 2014 from <http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf>.

Vreven, D., & McFadden, S. (2007). Success rates for students taking compressed and regular length development courses in the community college. *Community College Journal of Research and Practice*, 34(1-2), 39-54. doi: [10.1080/10668920903385806](https://doi.org/10.1080/10668920903385806).

Walvoord, B. E. (2008). *Teaching and learning in college introductory religion courses*. Malden, MA: Blackwell Publishing.

Webster, J. S., Runions, E., Gallagher, E. V., Lopez, D. C., McGinn, S. E., Penner, T. C., & Howell, D. B. (2012). Student learning outcomes for biblical studies in the liberal arts. *Teaching Theology & Religion*, 15, 262-83. doi: [10.1111/j.1467-9647.2012.00807.x](https://doi.org/10.1111/j.1467-9647.2012.00807.x).

Wlodkowski, R. (2003). Accelerated learning in colleges and universities. *New Directions for Adult and Continuing Education*, 97, 5-15. doi: [10.1002/ace.84](https://doi.org/10.1002/ace.84).

Wlodkowski, R. J., & Ginsberg, M. B. (2010). *Teaching intensive and accelerated courses: Instruction that motivates learning*. San Francisco: Jossey-Bass.