An Evaluation of the Effectiveness of WileyPLUS in Higher Education

EFFICACY STUDY RESULTS
Background

In the 2009-2010 academic year, The University of Tennessee’s Institute for Assessment and Evaluation (IAE) and J. Wiley & Sons, Inc. (Wiley) collaborated to conduct an evaluation of WileyPLUS on its several users (students, instructors, and academic departments). WileyPLUS is an online teaching and learning environment that integrates the digital textbook with resources for both instructors and students, allowing for a customizable learning experience catered to each student’s learning style.

This study identifies the effect that WileyPLUS has on students’ learning gains and shows that students who use WileyPLUS have a better learning outcome than do those who do not use the online learning environment.

IAE’s evaluation consisted of a multi-dimensional, multiple data source study with 22 case studies and document analysis; 11 efficacy studies exploring the knowledge gains, motivational changes, and course completion rates between students in course sections in which WileyPLUS was used and in comparable sections (same course) where WileyPLUS was not used; an inquiry into the reasons former users lapsed in their use of the WileyPLUS; and analysis of student and instructor survey data collected by Wiley in Fall 2009.

The results in this report represent Broadview Analytics’ analysis of the efficacy data gathered by IAE to evaluate the outcomes for students using WileyPLUS.

The Quasi-Experimental Design for Testing the Impact of WileyPLUS

<table>
<thead>
<tr>
<th>Students enroll in preferred course</th>
<th>Classes are randomly assigned to a group</th>
<th>The Semester Elapses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A</td>
<td>Pre-test</td>
<td>Test Group Results</td>
</tr>
<tr>
<td>Class B</td>
<td>Take the Course</td>
<td>Control Group Results</td>
</tr>
<tr>
<td></td>
<td>Post-test</td>
<td></td>
</tr>
</tbody>
</table>
Study Design

The study relied on a quasi-experimental design and analysis that identified and estimated the effect learning gain from using WileyPLUS. IAE recruited 11 Institutes of Higher Learning (IHLs) to participate in the efficacy study, including both 2-year and 4-year IHLs. In order to ensure that the results were not subject-specific, 7 different disciplines are represented across the various IHLs. Additionally, the type of delivery for the class (e.g., in-person, online, hybrid) was also included to account for the effect in these different environments. In total, 497 students participated in the study, with 267 in the test group and 230 in the control group.

Students at the participating institutions who wished to take one of the relevant courses had the opportunity to enroll in one of two seemingly identical sections. The enrolled students were not aware that the two sections were part of an experiment, where the instructor chose one section to be a control group and the other to be a test group. In the test group, the instructor led the course while utilizing WileyPLUS; the instructors in the control group did not use WileyPLUS. The effect of using WileyPLUS was then measured by the difference in learning outcomes between the groups, defined by the students’ scores on a post-course assessment test, which was designed and administered by their respective instructor.

To ensure that the effects of WileyPLUS were not confounded by any other factor, students were given an assessment test before the course began. This pre-test was then paired with the post-test in order to control for any differences in students’ ability prior to entering the class. This type of pre-post design is known as Analysis of Covariance, or ANCOVA.

The results of the pre-test from seven of the disciplines were quite different from the results of two of the disciplines: Engineering and Organic Chemistry. The results from the seven disciplines each show a normal...
distribution of pre-test scores, while Engineering and Organic Chemistry do not. Since these outliers were not consistent with the rest of the data, they were removed.

**Results**

**Student Pre- and Post-Class Performance**

As would be expected, students who completed a course in a subject became more knowledgeable about the subject. On average, students showed a substantial increase from their pre-test score (40.6) to their post-test score (68.1), representing a 68% improvement merely from participating in the course.

The effect was more pronounced among those who used WileyPLUS. For students who were in the test group, the improvement from the pre-test (38.3) to post-test (68.5) represents a 79% improvement (see Chart 1). But for those in the control group, the improvement from pre-test (44.5) to post-test (67.3) represents a 51% increase.

**The Effect of the Test Group versus the Control Group**

The size of the effect of using *WileyPLUS* is given by the ANCOVA model that controls for the pre-existing abilities of students. Adjusted for the pre-test score, WileyPLUS increases the final score from 65.5 to 69.8, a 7% increase in the outcome. This is a significant effect (degrees of freedom = 437; F-statistic = 5.6; p-value = 0.018). Students who use *WileyPLUS* improve their outcome on the post-test (see Chart 2).
The Effect at 2-year Institutions

The impact of using WileyPLUS is even more pronounced at two-year institutions. When using WileyPLUS, students improve their post-test result by 5.4 points (degrees of freedom = 199; F-statistic = 4.5; p-value = 0.034). This is an improvement of 9% in the outcome. (See Chart 3.) The average student at a 2-year school who uses WileyPLUS sees an improvement that is 28% better than the improvement among all students in general.

Classes that utilize WileyPLUS average a measured outcome that is 7% higher than those classes that do not.
Conclusion

Using WileyPLUS results in students receiving a better learning outcome*, on average, than if they did not use the system. This effect is even more pronounced for students at 2-year institutions.

Students participating in a course where WileyPLUS is used have a significantly better learning outcome at the end of the course. They increased their knowledge of the subject by 79%. More importantly, users of WileyPLUS performed better in the instructor-administered and designed post-test than did their counterparts who did not use WileyPLUS.

WileyPLUS helps the least-prepared students catch up to the rest of the students in the course. The gap between the least-prepared and average students is narrower after completing a course in which students chose to use WileyPLUS.

*Learning outcome refers, for the purposes of this report, to the performance of the students on an instructor-designed proxy for a final examination.

This is proprietary information of J. Wiley & Sons, Inc.
For more information about the Broadview Analytics impact analysis contained herein or The University of Tennessee’s Institute for Assessment and Evaluation study, please contact Petra Steriti at psteriti@wiley.com.

About The University of Tennessee’s Institute for Assessment and Evaluation
Established in 1994, the Institute for Assessment and Evaluation is an applied research and development community providing leadership in the field of education assessment and evaluation. Located within the College of Education, Health, and Human Sciences at the University of Tennessee’s Knoxville campus, the IAE group is dedicated to the development and implementation of effective strategies and technologies for assessing and evaluating educational organizations, instruction, and programs.

About Broadview Analytics
Broadview Analytics, Inc. is a full service quantitative-based consulting firm that assists clients in understanding complex markets and systems. By using comprehensive analytical processes, BA Inc. uncovers deep relationships that exist within a wide range of data – market research, operational or financial performance. Clients are supplied with a clear, actionable fact base that address core strategic business decisions. BA Inc. is lead by Joseph Stephan, a business executive with over 10 years of experience in market research, management consulting and finance.
Frequently Asked Questions

Study Design:

**Q: Within an institution, did the same instructor teach the course for the test group and the control group? If not, why?**

The efficacy study sites consisted of a one- or a two-instructor design. When recruiting for the study, IAE encountered a number of instructors who were unwilling to teach without *WileyPLUS*, thus a single-instructor design was not always possible, especially in light of other design criteria (e.g., a relatively balanced representation by school type (2-year and 4-year) and the inclusion of multiple disciplines.) Where a two-instructor design was used, instructors in the test and control groups were teaching the same course and administered the same pre- and post-test measurements for the study.

**Q: What was the range for the class sizes for the study sites?**

The average class size is 35, with a median of 29. The percent distribution was as follows: 25 or fewer students in the class 29%; 26-39 students 47%; 40 or more students 24%. For two year schools, the distribution is as follows: 25 or fewer students in the class 30%; 26-39 students 60%; 40 or more students 10%. Future research will address *WileyPLUS* efficacy in larger classes.

**Q: What were the pre- and post-course assessment tests?**

These were instructor designed assessments to ensure that the questions measured knowledge on topics that were relevant to each instructor’s course. Each instructor administered the same test at the start of the course and at the conclusion. Assessments were much like a typical final exam, though instructors did not necessarily use the score from the post-test as a student’s final exam grade.

Results:

**Q: What was the size of study, that is, how many different students were involved?**

There were 591 different students in the final sample. There were 377 in the test group and 214 in the control group.

**Q: Were the results evaluated by the letter grade that students earned to assess whether students earning certain grades performed better than others? If not, why?**

For this study, efficacy by student grades was not evaluated due to the study scope and student confidentiality. Based on pre- and post-assessment scores, the study revealed that the lowest performing students score better on the post-test as compared to the better performing students. Wiley hopes to evaluate *WileyPLUS*’ efficacy by student grades in future research.

**Q: Were students taking the course in hybrid and online settings included in the study? If not, why?**

The study did not evaluate the performance of students in hybrid and online...
settings. Because quasi-experimental studies are by their nature difficult to set up and execute, the scope of this study was carefully limited.

The United States Department of Education conducted a meta-analysis and review of online learning studies and in their May 2009 report titled *Evaluation of Evidence-Based Practices in Online Learning*, there is evidence to support improved outcomes by students taking hybrid and distance learning courses. For a copy of the report, visit the DOE website (http://www2.ed.gov/rschstat/eval/tech/evidence-based-practices/finalreport.pdf).

**Q:** Was the level of student and instructor use of WileyPLUS evaluated quantitatively? If so, what was the relationship to improvement?

A criterion for participation was that all instructors had to have at least two years of experience using WileyPLUS. Actual instructor and student use was evaluated qualitatively and based on instructor reporting. In the new version of WileyPLUS launching in Fall 2011, it is possible to measure student time on task and future research will incorporate this.

**Q:** What about the Organic Chemistry and Engineering classes made them different; and ultimately lead to their exclusion from the analysis?

The distribution of the pre-test scores for these two disciplines was categorically different from the pre-test scores in the other disciplines. The number of students in these two courses who scored very low on the pre-test was quite high. This lead to any effect of WileyPLUS use being dwarfed by the increase in discipline knowledge from simply taking the course.

**Q:** Are results available for the individual disciplines? Can the current analysis tell us anything about possible trends in the disciplines?

There are not enough students within any single discipline to test for a result solely within an individual discipline. Even though the structure of the experiment incorporates a time aspect with the pre-test and post-test methodology, it was not designed to analyze trends. Wiley is open to future longitudinal studies that track the effectiveness over multiple semesters.

The disciplines evaluated in the impact study were Accounting, Anatomy & Physiology, Biology, Chemistry, Geography, Mathematics and Physics. Case studies covered additional disciplines as follows: Engineering, Psychology and Spanish.

**Q:** How, if at all, can the successes demonstrated in this study be related to letter grades as the outcome?

The final letter grade of a student depends on many factors, including, but not limited to, the assessment methods that an instructor uses, the weighting for those and instructor discretion. Because of these many factors, it is not possible to say with certainty what an equivalent letter grade may be. However, depending on how an instructor grades, it is possible that use of WileyPLUS can improve student scores by as much as half or a whole letter grade on a final
exam. This is because pre-/post-assessment test, which was the proxy for measuring student learning gains, can be likened to a final exam.

**Q: Did the data show a normal distribution for post-test scores for the treatment group?**
The data for the post-test scores follows a normal distribution.

**Q: What, if anything, can this study suggest about the relationship between *WileyPLUS* use and course retention rates? What, if anything, can this study tell us about the relationship between *WileyPLUS* use and graduation rates?**
The current research sought to take a first step toward understanding efficacy as a whole due to constraints such as time and student privacy. The current study, which was conducted during a single term, measured retention in terms of a student having received a grade in the course. No differences were found, likely due to the small sample size of each individual class and because this the single time period (i.e., one term) in which the study was conducted. Wiley is evaluating the possibility of conducting future longitudinal research.

**Overall:**

**Q: What was John Wiley & Sons’ role in this study?**
Wiley was for the most part independent of this study. Wiley was the research sponsor, though company representatives took a hands-off approach to ensure that Wiley did not influence the research outcomes. All aspects of the research were managed and executed by IAE and some analysis and reporting was conducted by Broadview Analytics.

**Q: Where can I get further information on this study?**
IAE’s evaluation consisted of a multi-dimensional, multiple data source study with 22 case studies and document analysis; 11 efficacy studies exploring the knowledge gains, motivational changes, and course completion rates between students in course sections in which *WileyPLUS* was used and in comparable sections (same course) where *WileyPLUS* was not used; an inquiry into the reasons former users lapsed in their use of the *WileyPLUS*; and analysis of student and instructor survey data collected by Wiley in Fall 2009.

To obtain the case studies and an executive summary of IAE’s findings or the most recent results of the *WileyPLUS* student and instructor surveys, please contact your publisher’s representative at Wiley.