First, Do No Harm: Effective, Ineffective and Counterproductive Teaching Methods

Emmeline de Pillis
University of Hawaii at Hilo

Gene Johnson
University of Hawaii at Hilo

While excellent reviews of educational methods are available (e.g. Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Pashler et al., 2007; Rohrer & Pashler, 2010; Weinstein, McDermott, & Roediger, 2010), the contribution of the present paper is to distinguish between actions that can be taken by individual students or instructors (e.g. using flashcards) from those that require institutional action (e.g. reducing class size). In addition, we include both learning and completion effects where available. Our review of the literature found that small class sizes, repeated testing, and the use of full-time faculty are associated with improved outcomes.

INTRODUCTION

In response to public criticism of the quality of college education (e.g.Arum & Roksa, 2011; Bennett & Wilezol, 2013), the Association to Advance Collegiate Schools of Business - International (AACSB) and other accreditors have increased their emphasis on assessing student outcomes (Lynn Johnson, 2012; Koppel & Hollister, 2009; Kuh, Jankowski, Ikenberry, & Kinzie, 2014; Maki, 2012). Universities are expected to define what we are trying to teach, measure how well our students are learning these things, and "close the loop" by going back and adjusting curriculum to address any weaknesses. Presumably, measurable improvements in learning and student success will follow, completing a cycle in the continuous improvement model.

However, intuitively appealing educational interventions can be ineffective (Pashler, McDaniel, Rohrer, & Bjork, 2008) or even counterproductive (Forsyth, Lawrence, Burnette, & Baumeister, 2007), and if we are to make strides toward continuous improvement, we must first take care to consider the evidence in order to decide which measures are likely to bring about positive change.

Purpose—Do No Harm

Before committing resources to an intervention, it seems appropriate to examine the evidence in favor of that intervention. We borrow our working definition of evidence-based education from the literature on evidence-based management: The systematic use of the best available evidence to improve practice" (Reay, Berta, & Kohn, 2009, p. 5). "It seems reasonable to expect that those preparing students for society and the world of work would themselves take an evidence-based approach to their teaching." (Klimoski &

Amos, 2012, p. 686). An ill-considered decision can cause real harm, either by wasting resources on an unproven and ineffective method, or worse, by reducing student learning or likelihood of completion.

Finding evidence upon which to base recommendations can be less than straightforward. As in any field, one study's findings might conflict with another, or the generalizability of a set of findings might be in doubt (Kvernbekk, 2011). In addition, the field of education has not been uniformly welcoming of an evidence-based approach, with the result that many important questions may have gone untested (Cook, 2002; Cooper, Levin, & Campbell, 2009). The framework that follows will allow us to consider the quality of evidence in evaluating the effect of various interventions on student learning.

Procedure

To determine popular interventions at the postsecondary level, we started with the practice guide *Organizing Instruction and Study to Improve Student Learning*. (Pashler et al., 2007) This document is posted on the website of the Institute of Education Sciences *What Works Clearinghouse*. The guide recommends practices that have a demonstrable and positive effect on student learning. Research in K-12 learning is far more plentiful than research in postsecondary . The authors of the practice guide acknowledge that "Although the findings described here are probably as pertinent to college instruction as to lower grades, our most direct concern in producing this guide has been education from 3rd through 12th grade (Pashler et al., 2007, p. 3)" Recommendations supported by are: (1) Space learning over time; (2) Interleave worked example solutions and problem-solving exercises; (3) Combine graphics with verbal descriptions; (4) Connect and integrate abstract and concrete representations of concepts; (5) Use quizzing to promote learning; (6) Help students allocate study time (support for this was weak) and (7) Ask deep questions.

To round out our review of interventions covered in the What Works Clearinghouse, we added topics (such as classroom flipping) that are new and popular in the higher education press, and more perennial issues (such as class size) that have become contentious as appropriations to higher education have decreased.

Inclusion Criteria

Studies were included if they were published in scholarly journals and had postsecondary students as learners. We sought out studies that met the guidelines of the highest quality of evidence. For example, observational studies were excluded unless no relevant controlled study was available. To denote the quality of evidence, we used Reay and associates' six – level framework:

Level 1: Large scale randomized, controlled trials (RCTs) or meta-analyses

Level 2: Evidence from small sample RCTs, systematic literature reviews

Level 3: Retrospective case control studies, prospective cohort studies, multisite observational studies

Level 4: Small sample, single site observational studies

Level 5: Descriptive studies, case studies

Level 6: Expert opinion, anecdotal evidence. (2009, p. 9)

Refinement of Categories

With the introduction of college completion into the discussion, tensions have arisen between credential-granting and student learning (Friedel, Thornton, D'Amico, & Katsinas, 2013; Humphreys, 2012). Sufficient headcount is necessary for a school's financial health, and examining completion rates can highlight programmatic bottlenecks. A single-minded focus on retention and completion, however, can lead to a "student as customer" mindset (Laing & Laing, 2011), where challenging coursework is

watered down, cheating goes unpunished, and student learning suffers. For this reason, we examine effects on student learning and completion rates separately.

In addition, we distinguish classroom interventions (such as quizzes) and structural interventions (such as class size). The former may be employed by individual faculty members, while the latter are controlled by the institution.

Limitations of Available Evidence

While the randomized, controlled, double blind study provides the highest quality of evidence (D'Agostino & Kwan, 1995; Reay et al., 2009), generating such evidence on educational practices is not as straightforward as one might hope. Randomization in particular can present ethical and operational difficulties. For example, studies of online learning can be thwarted by student participants dropping their randomly-assigned sections and re-enrolling in the sections they prefer. Other types of studies should not be completely ignored, and may be useful if they are well controlled and interpreted with caution (D'Agostino & Kwan, 1995; Slavin, 2008).

RESULTS

Institutional Level

Intervention	Effect on Learning	Effect on Completion
Limiting class size	Positive. Increased class size is associated with decreased student academic performance. The optimal class size appears to be between10 and 15, while the point at which learning begins to deteriorate is between 10 and 25. The point of no return, beyond which increasing the number of students does little additional harm to learning, is in the range of26 to 100 (Bandiera, Larcinese, & Rasul, 2010; C. M. Campbell, Jimenez, & Cruz Paul, 2013; Cuseo, 2007; De Paola & Scoppa, 2011; I. Johnson, 2010; Kokkelenberg, Dillon, & Christy, 2008). Level: 3	Positive. Increases in student-faculty ratios account for over three-quarters of the decrease in completion rates relative to the 1970s (Bound, Lovenheim, & Turner, 2010). Level: 3
Using full time faculty	Mostly positive. Contingent faculty on average spend significantly less time preparing for class and advising students (Umbach, 2008) and assign significantly higher grades (I. Johnson, 2011), which may reflect a more forgiving grading scale. Contingent instructors typically rely on high student evaluations for contract renewal, and may not have departmental support to hold	Positive. As exposure to part-time faculty increases, retention (Eagan & Jaeger, 2008; Jaeger & Eagan, 2011), graduation rates (Ehrenberg & Zhang, 2005) and likelihood of transfer (Kevin Eagan & Jaeger, 2009) decrease. However one study found no significant in the field of business (E. P. Bettinger & Long, 2010). This may reflect the type of adjunct; business schools often use working professionals. Other departments

	students to high standards. This is not to put any blame on contingent faculty; rather, the institutional support and incentives for contingent faculty are different from those for full time faculty. One recent study at Northwestern, however, found that students of non tenure track instructors enjoyed improved learning outcomes (Figlio, Schapiro, & Soter, 2013). Level: 3	often rely on "freeway flyers," who are patching together several part-time jobs at various institutions. Level: 3
Intrusive advising/coaching	Positive. In one study, students required to meet with academic mentors achieved higher grades in the same classes compared to students who were not assigned mentors (Sandner, 2013). Level:3.	Positive. A randomized trial of the coaching services provided by InsideTrack showed that completion rates were four percent higher in the treatment group. (E. Bettinger & Baker, 2011) Level:2
First year student success programs	Ineffective. The higher grades observed among participants in first year success courses (e.g.Cho & Karp, 2012) appear to be due to self-selection variables (Pike, Hansen, & Lin, 2011) Level:3	Ineffective. A study of the Opening Doors program, using random assignment, found that the program did not improve students' graduation rates. Four years after the start of the study, around seven percent of both the program and control group students had earned a degree or a certificate (Weiss, Brock, Sommo, Rudd, & Turner, 2011). Level: 1
Supplementing with intelligent tutoring systems	Promising. Experiments in intelligent tutoring systems have shown significant improvements in learning (Ghee Ming, Chai, & Maskell, 2010; Philippe, 2013). Level:2	Unknown, but judicious use of intelligent tutoring systems might free up instructors for more numerous and smaller class sections.
Flipping the classroom. A flipped classroom is one where students absorb the basic material outside of class, and then do an active assignment under the instructor's guidance.	Mixed to ineffective. Enthusiastic commentary abounds (e.g.Bergmann, 2012; Berrett, 2012), but there is little randomized research that compares learning outcomes (Bishop, 2013). Studies can be confounded by other variables such as simultaneous tightening of admissions standards (e.g. Moravec, Williams, Aguilar-Roca, & O'Dowd, 2010). One dissertation study found	Unknown.

	worse grades and less confidence in a "flipped" math class compared to a traditional course (Strayer, 2007). Another showed no clear benefit of a flipped classroom, and increased off-task behavior(L Johnson & Renner, 2012). Researchers at Harvey Mudd College are currently working on a more rigorous and nuanced study of flipping, including determining the conditions under which flipping is more or less effective (Lape, Levy, & Yong, 2014). Level:2	
Moving full-semester classes online	Negative. Despite a Department of Education report favorable to online courses (U.S. Department of Education, 2010), the effects of putting full semester courses online are negative (Jaggars & Bailey, 2010). A multisite observational study (Level 3) found a "robust negative impact of online course taking" for both English and math courses (Xu & Jaggars, 2011, p. 360). Level: 2	Negative. A review of the literature indicates that "online coursework—at least as it is currently and typically implemented—may hinder progression for low-income and underprepared students."(Jaggars, 2011, p. 2) Level:2

Instructor Level

The interventions below can be employed by individual instructors, independent of institutional policy.

Intervention	Effect on Learning	Effect on Completion
Spacing out learning over time.	Positive. Inserting a time interval between quizzes or practice sessions on the same material has long been known to have a beneficial effect on learning (Ebbinghaus, Ruger, & Bussenius, 1913). The "spacing effect" appears to increase both the amount of material learned and the length of time it is retained. This recommendation is included in the Practice Guide. (Dunlosky et al., 2013; Pashler et al., 2008; Pashler, Zarow, & Triplett, 2003; Pavlik & Anderson, 2008; Taylor & Rohrer, 2010). Level:2	Unknown
Interleaving reading with working problems.	Positive. Alternating between reading worked solutions and working out solutions on one's own has been shown to be effective in mathematics and science classes. This recommendation is included in the Practice	Unknown

	Guide. Level: 2	
Combining graphics with verbal descriptions	Positive. This has been shown to be effective for mathematics and science instruction. This recommendation is included in the Practice Guide. Level: 2	
Integrating abstract and concrete representations	Positive. While students may appreciate concrete examples, without the underlying theory students have trouble generalizing an idea. Both abstract and concrete examples help students understand and generalize. This has been tested mainly in mathematics instruction. (De Bock, Deprez, Van Dooren, Roelens, & Verschaffel, 2011; Kaminski, Sloutsky, & Heckler, 2008). This recommendation is included in the Practice Guide. Level: 2	
Incorporating frequent retrieval tasks such as quizzes	Positive. Retrieving information from memory with quizzes or flashcards, rather than just studying or rereading material, improves retention of material. (Butler, 2010; Dunlosky et al., 2013; Karpicke & Blunt, 2011; Pashler et al., 2008; Roediger & Karpicke, 2006; Vaughn & Rawson, 2011; Weinstein et al., 2010). This recommendation is included in the Practice Guide. Level: 2	Unknown
Introducing "deep" questions	Positive. In the Practice Guide, the authors note that the evidence supporting this practice is strongly supported and applicable over a wide range of subjects. "Deep" questions are of the type "why, why-not, how, what-if, how does X compare to Y, and what is the evidence for X?" (Pashler et al., 2007, p. 29) Level: 1	
Teaching to students' unique learning style	Ineffective. A systematic review of studies attempting to validate the learning styles approach found little to no evidence supporting this approach. There is no evidence that teaching a given subject using an individual's preferred style results in improved learning. (Pashler et al., 2008). Level:2	Unknown. There is no evidence of positive or negative effect on completion.
Bolstering student self-esteem	Negative. Compared to a similar control group, D and F students who received self-esteem-building messages had significantly decreased academic performance (Forsyth et al., 2007). Level: 2	Unknown, but given the negative effect on learning, it is unlikely to have a positive effect on completion.

RECOMMENDATIONS

Learning Research is Not for Amateurs

It is important to remember that homegrown assessment is not "level one" research. Before deciding that something works in the classroom because it seems right or feels good, it is imperative to search for valid studies. Faculty members are specialists in their own fields, but they do not necessarily have

expertise in psychometrics or assessment. Thus, results derived from homegrown assessment efforts will generally fall quite low on the scale of rigor of evidence.

Assessment may be viewed as similar to the diagnostic work physicians perform (Klimoski & Amos, 2012). The evidence based approach to medicine requires the physician to seek the best cure (intervention) from all of the evidence available, and to then use judgment and experience decide whether and how to implement it. Practicing physicians are not necessarily medical researchers, but they are expected to keep up with current research in their respective fields. Similarly, an evidence based approach to continuous improvement would require administrators and educators to seek the best way to improve student learning (intervention) from all of the evidence available. Rarely would that intervention be identified and supported by homegrown assessment activities alone.

Use a Value Added Approach

Less-selective institutions will always find themselves at a disadvantage in measures of student and alumni outcomes unless incoming student ability is taken into account (Higher Education Research Inst, 2003). The "value added" approach appears to be gaining momentum (e.g.HCM Strategists, 2012; Liu, 2011; Thomas, 2010). This approach insures that institutions are not punished for having inclusive admissions policies.

The relationship between cognitive ability and academic performance is strong and intractable. The SAT, widely used for college admissions, predicts both grades and later career outcomes (DeAngelo, Franke, Hurtado, Pryor, & Tran, 2011). The SAT and similar tests are closely correlated with other measures of cognitive aptitude (Roediger & Karpicke, 2006), which remains the single best predictor of academic performance (Hambrick & Meinz, 2011; Kaufman, Reynolds, Liu, Kaufman, & McGrew, 2012).

Attempts to level the playing field through remedial or developmental education have had very limited success. Randomized studies show little or no measurable effect on performance (e.g.Wagner, 2011), or at best, small gains for a high cost (Melguizo, Bos, & Prather, 2011). Remediation "might promote early persistence in college, but it does not necessarily help students on the margin of passing the placement cutoff make long-term progress toward earning a degree." (Calcagno & Long, 2008, p. 1) A large scale observational study concluded that remedial education, despite its high cost, improved neither academic nor labor market outcomes (Martorell & McFarlin Jr, 2011).

The optimal time for remediation appears to be in early childhood (F. A. Campbell et al., 2012; Muennig et al., 2011). The gap between children of affluent families and those of poor families is measurable well before age 2 (Taylor & Rohrer, 2010), indicating that efforts to catch up the children of less affluent families may need to begin in infancy; college remediation efforts, however well-intentioned, come far too late.

Conclusion

An evidence based approach to closing the loop requires us to study the relevant literature in search of rigorous evidence identifying and supporting such interventions, and to base our decisions on that evidence. Indeed, the literature does support a number of interventions that are good candidates for improving student learning.

The use of an evidence based approach should help to minimize inappropriate and/or ineffective interventions by reducing the likelihood of changes being made based solely on single-entity assessment data (which does not rank very high in quality of evidence) and by leading educators toward a more scientific approach. Perhaps one of the most damaging forces in education today is the pervasive pressure to "just change something!" when in fact the most appropriate position may very well be "first, do no harm."

REFERENCES

Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses.* Chicago: University of Chicago Press.

Bandiera, O., Larcinese, V., & Rasul, I. (2010). Heterogeneous Class Size Effects: New Evidence from a Panel of University Students*. *The Economic Journal, 120*(549), 1365-1398. doi: 10.1111/j.1468-0297.2010.02364.x

Bennett, W. J., & Wilezol, D. (2013). *Is college worth it?: A former United States Secretary of Education and a liberal arts graduate expose the broken promise of higher education*: Thomas Nelson Publishers.

Bergmann, J. (2012). Flip Your Classroom: Talk To Every Student In Every Class Every Day Author: Jonathan Bergmann, Aaron Sams, Publisher: Inte.

Berrett, D. (2012). How'flipping'the classroom can improve the traditional lecture. *The chronicle of higher education, 12*.

Bettinger, E., & Baker, R. (2011). The effects of student coaching in college: An evaluation of a randomized experiment in student mentoring: National Bureau of Economic Research.

Bettinger, E. P., & Long, B. T. (2010). Does cheaper mean better? The impact of using adjunct instructors on student outcomes. *Review of Economics & Statistics, 92*(3), 598-613.

Bishop, J. L. (2013). The Flipped Classroom: A Survey of the Research. *Proceedings of the 2013 American Society for Engineering Education Annual Conference*.

Bound, J., Lovenheim, M. F., & Turner, S. (2010). Why Have College Completion Rates Declined? An Analysis of Changing Student Preparation and Collegiate Resources. *American Economic Journal: Applied Economics, 2*(3), 129-157. doi: doi: 10.1257/app.2.3.129

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(5), 1118-1133. doi: 10.1037/a0019902

Campbell, C. M., Jimenez, M., & Cruz Paul, T. (2013). COLLEGE EDUCATIONAL QUALITY (CEQ) PROJECT 2013 Pilot Study 1 *COLLEGE EDUCATIONAL QUALITY (CEQ) PROJECT*. New York, New York: Teachers College, Columbia University.

Campbell, F. A., Pungello, E. P., Burchinal, M., Kainz, K., Pan, Y., Wasik, B. H., . . . Ramey, C. T. (2012). Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. *Developmental Psychology*, No Pagination Specified. doi: 10.1037/a0026644

Cho, S.-W., & Karp, M. M. (2012). Student Success Courses and Educational Outcomes at Virginia Community Colleges. In C. C. R. Center (Ed.), *CCRC Assessment of Evidence Series* (pp. 1-19). New York, NY: Community College Research Center.

Cook, T. D. (2002). Randomized Experiments in Educational Policy Research: A Critical Examination of the Reasons the Educational Evaluation Community has Offered for not Doing Them. *Educational Evaluation and Policy Analysis, 24*(3), 175-199. doi: 10.3102/01623737024003175

Cooper, A., Levin, B., & Campbell, C. (2009). The growing (but still limited) importance of evidence in education policy and practice. *Journal of Educational Change, 10*(2-3), 159-171. doi: 10.1007/s10833-009-9107-0

Cuseo, J. (2007). The empirical case against large class size: adverse effects on the teaching, learning, and retention of first-year students. *The Journal of Faculty Development, 21*(1), 5-21.

D'Agostino, R. B., & Kwan, H. (1995). Measuring effectiveness. What to expect without a randomized control group. *Medical Care, 33*(4 Suppl), AS95-AS105.

De Bock, D., Deprez, J., Van Dooren, W., Roelens, M., & Verschaffel, L. (2011). Abstract or concrete examples in learning mathematics? A replication and elaboration of Kaminski, Sloutsky, and Heckler's study. *Journal for research in Mathematics Education, 42*(2), 109-126.

De Paola, M., & Scoppa, V. (2011). The effects of class size on the achievement of college students *The Manchester School, 79*(6), 1061-1079. doi: 10.1111/j.1467-9957.2010.02208.x

DeAngelo, L., Franke, R., Hurtado, S., Pryor, J. H., & Tran, S. (2011). Completing College: Assessing graduation rates at four-year institutions. Los Angeles: Higher Education Research Institute, UCLA.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest, 14*(1), 4-58. doi: 10.1177/1529100612453266

Eagan, M. K., & Jaeger, A. J. (2008). Closing the gate: Part-Time faculty instruction in gatekeeper courses and first-year persistence. *New Directions for Teaching and Learning, 2008*(115), 39-53. doi: 10.1002/tl.324

Ebbinghaus, H., Ruger, H. A., & Bussenius, C. E. (1913). Retention and obliviscence as a function of the time *Memory: A contribution to experimental psychology* (pp. 62-80). New York, NY, US: Teachers College Press.

Ehrenberg, R. G., & Zhang, L. (2005). Do Tenured and Tenure-Track Faculty Matter? *Journal of Human Resources, XL*(3), 647-659. doi: 10.3368/jhr.XL.3.647

Figlio, D. N., Schapiro, M. O., & Soter, K. B. (2013). Are Tenure Track Professors Better Teachers? : National Bureau of Economic Research.

Forsyth, D. R., Lawrence, N. K., Burnette, J. L., & Baumeister, R. F. (2007). Attempting to improve the academic performance of struggling college students by bolstering their self-esteem: an intervention that backfired. *Journal of Social & Clinical Psychology, 26*(4), 447-459.

Friedel, J. N., Thornton, Z. M., D'Amico, M. M., & Katsinas, S. G. (2013). Performance-Based Funding: The National Landscape. In U. o. A. E. P. Center (Ed.), (pp. 20): University of Alabama.

Ghee Ming, G., Chai, Q., & Maskell, D. L. (2010). EpiList II: Closing the Loop in the Development of Generic Cognitive Skills. *IEEE Transactions on Systems, Man & Cybernetics: Part A, 40*(4), 676-685. doi: 10.1109/tsmca.2010.2041226

Hambrick, D. Z., & Meinz, E. J. (2011). Limits on the Predictive Power of Domain-Specific Experience and Knowledge in Skilled Performance. *Current Directions in Psychological Science, 20*(5), 275-279. doi: 10.1177/0963721411422061

HCM Strategists. (2012). Context for Success: Measuring Colleges' Impact. 2012, from http://www.hcmstrategists.com/contextforsuccess/papers.html

Higher Education Research Inst, I. L. A. C. A. (2003). *How "Good" Is Your Retention Rate? Using the CIRP Survey To Evaluate Undergraduate Persistence*.

Humphreys, D. (2012). What's wrong with the completion agenda—and what we can do about it. *Liberal Education, 98*(1), 158.

Jaeger, A. J., & Eagan, M. K. (2011). Examining Retention and Contingent Faculty Use in a State System of Public Higher Education. *Educational Policy, 25*(3), 507-537. doi: 10.1177/0895904810361723

Jaggars, S. S. (2011). Online Learning:Does It Help Low-Income and Underprepared Students? . In C. C. R. Center (Ed.), *CCRC Assessment of Evidence Series* (pp. 1-57). New York, NY: Community College Research Center.

Jaggars, S. S., & Bailey, T. (2010). Effectiveness of Fully Online Courses for College Students: Response to a Department of Education Meta-Analysis
In C. C. R. Center (Ed.), *CCRC Assessment of Evidence Series* (pp. 1-18). New York, NY: Community College Research Center.

Johnson, I. (2010). Class Size and Student Performance at a Public Research University: A Cross-Classified Model. *Research in Higher Education, 51*(8), 701-723. doi: 10.1007/s11162-010-9179-y

Johnson, I. (2011). Contingent Instructors and Student Outcomes: An Artifact or a Fact? *Research in Higher Education, 52*(8), 761-785. doi: 10.1007/s11162-011-9219-2

Johnson, L. (2012). Closing The Loop: Using Assessment Results To Modify The Curriculum So That Student Quantitative Reasoning Skills Are Enhanced. *American Journal of Business Education (AJBE), 5*(4), 465-468.

Johnson, L., & Renner, J. (2012). *Effect of the flipped classroom model on secondary computer applications course: student and teacher perceptions, questions and student achievement.* (Doctor of Education Dissertation), University of Louisville, Louisville, KY.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). Learning theory: The advantage of abstract examples in learning math. *Science, 320*(5875), 454-455.

Karpicke, J. D., & Blunt, J. R. (2011). Retrieval Practice Produces More Learning than Elaborative Studying with Concept Mapping. *Science, 331*(6018), 772-775. doi: 10.1126/science.1199327

Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive g and academic achievement g one and the same g? An exploration on the Woodcock–Johnson and Kaufman tests. *Intelligence, 40*(2), 123-138. doi: http://dx.doi.org/10.1016/j.intell.2012.01.009

Kevin Eagan, M., & Jaeger, A. (2009). Effects of Exposure to Part-time Faculty on Community College Transfer. *Research in Higher Education, 50*(2), 168-188. doi: 10.1007/s11162-008-9113-8

Klimoski, R., & Amos, B. (2012). Practicing Evidence-Based Education in Leadership Development. *Academy of Management Learning & Education, 11*(4), 685-702. doi: 10.5465/amle.2012.0018

Kokkelenberg, E. C., Dillon, M., & Christy, S. M. (2008). The effects of class size on student grades at a public university. *Economics of Education Review, 27*(2), 221-233. doi: 10.1016/j.econedurev.2006.09.011

Koppel, N. B., & Hollister, K. K. (2009). Retention Assessment Of Core Operations Management Topics For Business Administration Students *American Journal of Business Education, 2*(2), 31-38.

Kuh, G. D., Jankowski, N., Ikenberry, S. O., & Kinzie, J. (2014). Knowing what students know and can do: The current state of student learning outcomes assessment in US colleges and universities. *Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).*

Kvernbekk, T. (2011). The concept of evidence in evidence-based practice *Educational Theory, 61*(5), 515-532. doi: 10.1111/j.1741-5446.2011.00418.x

Laing, L., & Laing, G. K. (2011). The student as customer model and its impact on the academic leadership role in higher education. *Meeting the Challenges: Proceedings ofthe ATN Assessment Conference, 2011*, 117-123.

Lape, N., Levy, R., & Yong, D. (2014). Can flipped classrooms help students learn? We're trying to find out. *Slate*.

Liu, O. L. (2011). Value-added assessment in higher education: a comparison of two methods. *Higher Education, 61*(4), 445-461.

Maki, P. (2012). *Coming to Terms with Student Outcomes Assessment: Faculty and Administrators' Journeys to Integrating Assessment in Their Work and Institutional Culture*: Stylus Publishing (VA).

Melguizo, T., Bos, J., & Prather, G. (2011). Is Developmental Education Helping Community College Students Persist? A Critical Review of the Literature. *American Behavioral Scientist, 55*(2), 173-184.

Moravec, M., Williams, A., Aguilar-Roca, N., & O'Dowd, D. (2010). Learn before lecture: a strategy that improves learning outcomes in a large introductory biology class. . *CBE Life Sci Educ, 9*, 473–481.

Muennig, P., Robertson, D., Johnson, G., Campbell, F., Pungello, E. P., & Neidell, M. (2011). The effect of an early education program on adult health: The Carolina Abecedarian Project randomized controlled trial. *American Journal of Public Health, 101*(3), 512-516. doi: 10.2105/ajph.2010.200063

Pashler, H., Bain, P., Bottge, B., Graesser, A., Koedinger, K., McDaniel, M., & Metcalfe, J. (2007). Organizing Instruction and Study to Improve Student Learning. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning Styles. *Psychological Science in the Public Interest, 9*(3), 105-119. doi: 10.1111/j.1539-6053.2009.01038.x

Pashler, H., Zarow, G., & Triplett, B. (2003). Is Temporal Spacing of Tests Helpful Even When It Inflates Error Rates? *Journal of Experimental Psychology. Learning, Memory & Cognition, 29*(6), 1051-1057. doi: 10. w37/0278-7393.29.6.1051

Pavlik, P. I., & Anderson, J. R. (2008). Using a model to compute the optimal schedule of practice. *Journal of Experimental Psychology: Applied, 14*(2), 101-117. doi: 10.1037/1076-898x.14.2.101

Philippe, F.-V. (2013). A Multiparadigm Intelligent Tutoring System for Robotic Arm Training. *IEEE Transactions on Learning Technologies, 6*(4), 364-377.

Pike, G., Hansen, M., & Lin, C.-H. (2011). Using Instrumental Variables to Account for Selection Effects in Research on First-Year Programs. *Research in Higher Education, 52*(2), 194-214. doi: 10.1007/s11162-010-9188-x

Reay, T., Berta, W., & Kohn, M. K. (2009). What's the Evidence on Evidence-Based Management? *Academy of Management Perspectives, 23*(4), 5-18. doi: 10.5465/amp.2009.45590137

Roediger, H. L., & Karpicke, J. D. (2006). Test-Enhanced Learning. *Psychological Science, 17*(3), 249-255. doi: 10.1111/j.1467-9280.2006.01693.x

Rohrer, D., & Pashler, H. (2010). Recent Research on Human Learning Challenges Conventional Instructional Strategies. *Educational Researcher, 39*(5), 406-412. doi: 10.3102/0013189x10374770

Sandner, M. (2013). Quasi-Experimental Evaluation of a Student Mentoring Program: Diskussionspapiere der Wirtschaftswissenschaftlichen Fakultät, Universität Hannover.

Slavin, R. E. (2008). Perspectives on Evidence-Based Research in Education—What Works? Issues in Synthesizing Educational Program Evaluations. *Educational Researcher, 37*(1), 5-14. doi: 10.3102/0013189x08314117

Strayer, J. (2007). *The effects of the classroom flip on the learning environment: a comparison of learning activity in a traditional classroom and a flip classroom that used an intelligent tutoring system* (Doctoral Dissertation), The Ohio State University. Retrieved from http://etd.ohiolink.edu/send-pdf.cgi/Strayer%20Jeremy.pdf?osu1189523914

Taylor, K., & Rohrer, D. (2010). The Effects of Interleaved Practice. *Applied Cognitive Psychology, 24*, 837-848.

Thomas, S. M. (2010). Assessment and the Evaluation of Institutional Effectiveness. In P. Editors-in-Chief: Penelope, B. Eva, E. B. Barry McGawA2 - Editors-in-Chief: Penelope Peterson & M. Barry (Eds.), *International Encyclopedia of Education (Third Edition)* (pp. 172-180). Oxford: Elsevier.

U.S. Department of Education. (2010). Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies Washington, D.C.: Office of Planning, Evaluation, and Policy Development.

Umbach, P. D. (2008). *The effects of part-time faculty appointments on instructional techniques and commitment to teaching* Paper presented at the 33rd Annual Conference of the Association for the Study of Higher Education,, Jacksonville, FL.

Vaughn, K. E., & Rawson, K. A. (2011). Diagnosing Criterion-Level Effects on Memory. *Psychological Science, 22*(9), 1127-1131. doi: 10.1177/0956797611417724

Wagner, K. E. (2011). *Improving Student-Athletes' Writing Skills: Examining the Effects of Self-Regulated Strategy Development Coupled with Modified Reciprocal Teaching.* (M.A.), University of Oregon.

Weinstein, Y., McDermott, K. B., & Roediger, H. L., III. (2010). A Comparison of Study Strategies for Passages: Rereading, Answering Questions, and Generating Questions. *Journal of Experimental Psychology: Applied, 16*(3), 308-316.

Weiss, M., Brock, T., Sommo, C., Rudd, T., & Turner, M. C. (2011). Opening Doors: Serving Community College Students on Probation: MDRC.

Xu, D., & Jaggars, S. S. (2011). The Effectiveness of Distance Education Across Virginia's Community Colleges: Evidence From Introductory College-Level Math and English Courses. *Educational Evaluation and Policy Analysis, 33*(3), 360-377. doi: 10.3102/0162373711413814