Scholarship of Teaching and Learning in Psychology
2015, Vol. 1, No. 1, 79–91

2332-2101/15/$12.00 http://dx.doi.org/10.1037/stl0000021

TEACHER-READY RESEARCH REVIEW

The Applicability of Visible Learning to Higher Education

John Hattie
University of Melbourne

The Visible Learning research is based on a synthesis of 1200 meta-analyses relating to influences on achievement. This article focuses specifically on the evidence and implications for higher education teachers. As nearly every intervention can show some evidence of success, we need to ask not "What works?" but "What works best" and seek comparisons between different ways of influencing student learning. The major implications relate to teachers who work with others to seek evidence of their impact on students, who inform students early what success looks like especially about surface and deep learning, who provide appropriate levels of challenge and feedback, and who have aligned their claims about success, assessment, and teaching.

Keywords: learning, meta-analysis, teaching strategies

Today's university students are diverse, not necessarily self-regulated, having varying skills in learning strategies, and need to be deliberately taught. This begs for a robust discipline about the scholarship of teaching and learning at the university level to best identify what works. Whereas many thousands of studies on university learning exist, few major syntheses exist (see Pascarella & Terenzini, 2005). This article uses a synthesis of meta-analyses relating to university students from the Visible Learning (VL) research based on more than 65,000 studies including [1/4] million students aged 4 to 20–25 (Hattie, 2009, 2012; Hattie & Anderman, 2013; Hattie & Yates, 2014; Hattie, Masters, & Birch, 2015). The aim is to summarize key findings related to university aged cohort and discuss key implications for higher education (see also Hattie, 2011).

More is demanded from higher education students than knowing much and surviving through three-plus years of study. Forty-five years ago, Chickering (1969) outlined seven major outcomes from university study which are still applicable: achieving competence; managing emotions (including those that interfere with learning such as anger, anxiety, hopelessness and those that enable learning such as optimism, hopefulness); mature interpersonal relations (such respecting differences and working with peers); moving from autonomy to independence (including moving from needing assurance and approval of others to self-sufficiency), problem solving, and making decisions; establishing identity including enhanced self-esteem and self-efficacy; developing purpose (from Who am I? and Where am I? to Where am I going?); and developing integrity. Higher education is as much as about identity, reputational enhancement, and growing as it is about becoming knowledgeable, critics, and problem solvers. The (valuable) by-products are knowing more about a topic, being passionate about content, and being learned about a subject. With the more recent push for developing 21st century skills, there is a renewed emphasis on the latter 6, although developing subject matter expertise remains dominant.

The VL story argues that when teachers see teaching and learning through the eyes of their students, and when students become their own teachers then outcomes and engagement are maximized. It is not sufficient to stand up front and look and behold a great multitude from every nation and all tribes and peoples and

Correspondence concerning this article should be addressed to John Hattie, Faculty of Education, Melbourne Education Research Institute, University of Melbourne, Carlton, Victoria, Australia 3010. E-mail: jhattie@unimelb.edu.au

tongues, waxing lyrical about one's pet subject, and crying out loud that there is knowledge to be learnt, salvation for those who listen and who repeat after me, and it will finish on the hour. To be successful, university teachers need to think of themselves as evaluators and ask about the merit, worth, and significance of the impact of their interventions—essentially, successful educators actively practice the Scholarship of Teaching and Learning (SoTL).

This focus on impact demands questioning what "impact" means, as certainly a major finding from the VL research is that too often the answer is that surface not deep learning is maximized. Surface learning privileges knowing facts, ideas, and content, whereas deeper learning privileges knowing relations and connections between ideas and extending these ideas to other contexts. Surface is privileged, even noting that many university teachers proclaim to their students that this course is about understanding, making relations, and extending your knowledge; but students see that the assessments in the course value knowing much, and repeating back the major claims by the textbook or the instructor. Students are very strategic in working out what the teachers really values as opposed to what they might say they value. Hence, a very good place to start reviews of university courses is to analyze and review the skills needed by the student to successfully answer the assignments and tests. Such a review can be a powerful way to understand what impact looks like through the "eyes of students."

The evidence for the claims in this article comes from a synthesis of more than 1200 meta-analyses (Hattie, 2009, 2012). The aim of the synthesis was to place the various influences on student achievement along an underlying achievement continuum. As noted above, there are other outcomes besides achievement and there is a team based in Germany working on a similar synthesis for motivation and emotional outcomes; our team is completing a similar synthesis of learning strategies.

Method

The underlying metric used in the VL synthesis of meta-analyses is effect-size—which are based on the difference between implementing two methods (e.g., implementing a new method compared with other classes without this method) or the change from a pre to a post measure. These differences are standardized by using an estimate of the pooled variance across all the outcomes measurements (for how to estimate effect-sizes see Borenstein et al., 2009; Hedges & Olkin, 1985; Lipsey & Wilson, 2001). An effect size of $d = 1.0$, for example, indicates an increase of one standard deviation on the outcome, and is typically associated with advancing student's achievement by two to three years, improving the rate of learning by 50%, or a correlation between some influence and achievement of approximately $r = .50$. When implementing a new program, an effect size of 1.0 would mean that, on average, students receiving that treatment would exceed 84% of students not receiving that treatment. An effect-size less than .2 can be considered small, .4 average, and greater than .6 large (although care is needed given that there may be important moderators and mediators of any overall effect-sizes).

The data for the study involve 1200 meta-analyses relating to influences on student achievement, which encompass more than 65,000 studies, 150,000 effect sizes, and about [1/4] billion students. These studies related to the influence of some program, policy, or innovation on academic achievement in an education setting (early childhood, elementary, high, and tertiary). Although it is the case that most are derived from the K-12 sector, there are still many from the postsecondary sector, and the argument of this article is that the underlying messages underlying successful innovations are quite similar across the sectors.

Results and Discussion

Figure 1 shows the overall distribution of all the effect sizes from each of the 1200+ meta-analyses. The *y* axis represents the number of effects in each category, and the *x* axis gives the magnitude of effect sizes. Table 1 provides an overview of the 196 influences—the interest, however, is less in these rankings but in the story underlying why those in the top half differ from those in the bottom half—and that is the essence of the Visible Learning story (Hattie, 2009).

One of the surprising findings was that across the interventions that are commonly claimed to enhance student learning—nearly all of them

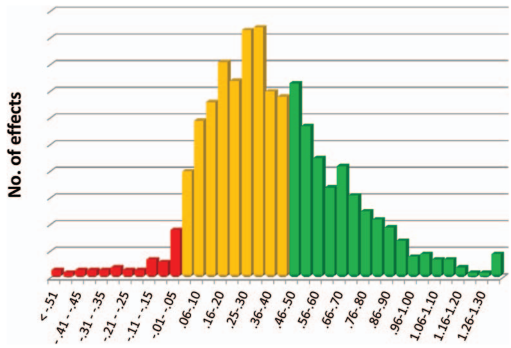

Figure 1. Distribution of effect sizes across all meta-analyses. See the online article for the color version of this figure.

have a positive impact on student learning. That is, almost everything works! But herein lies the greatest problem in education—every method seems to work relative to not implementing that method. This leads to many concluding that their particular method of teaching does enhance learning. But this claim can be made for almost every method. In many senses this is asking akin to the null hypothesis in research—does implementing my method mean an improvement (compared to “no” change) and the answer is yes. What we should be asking, instead, is the magnitude of the learning improvement—and this is where the major messages become obvious. Some interventions have dramatically higher impacts than others. The interest in VL is the underlying story about those influences higher compared with those lower than the average $d = .40$.

As an example of the preponderance of positive effects, Tomcho and Foels (2008) completed a meta-analysis of 197 studies relating to teaching activities and methods at the university level. Only 10 effects were negative (and they were close to zero), and everything else showed an achievement gain. The overall effect was .57 for knowledge, .27 for attitude, .39 for assessed behavior or skills, and .31 on grade outcomes. They cautioned that the sample may be biased toward those with caring, willing, and open dispositions to enhance their teaching (although these are the very skills underlying good teaching)—even so, such findings of positiveness are endemic in education.

The major VL story can be summarized by six key findings. impact on student learning is heightened:

1. When teachers believe their major role is to evaluate their impact ($d = .91$);
2. When teachers work together to know and evaluate their impact ($d = .91$);
3. When teachers base their teaching on students’ prior learning (what they bring to the lesson; $d = .85$);
4. When teachers explicitly inform the students about what success looks like near the start of a series of lessons (.77);
5. When teachers implement programs that have the optimal proportions of surface and deep learning ($d = .71$); and
6. When teachers set appropriate levels of challenge and never expect ‘do your best’ ($d = .57$).

Such impact is not only measured in changes in test scores; although test scores can be feedback to teachers about their impact, but also by the levels of investment, mastery, tolerance of dealing with mistakes, degrees of collaboration with others wishing to learn more, keenness to be successful, and high interest that students evoke.

The Visible Learning model asks educators to evaluate the quality of the evidence they can provide relating to these six key findings. A successful strategy to undertake this evaluation is to (a) strategically plan and self-review the teaching sessions to as to incorporate feedback about interventions and teaching impact; (b) ensure the presence of deliberate strategies for raising teaching impact; (c) use of student voice as part of the responses to interventions (i.e., listen to how students are understanding the teaching); and (d) gather and analyze data about impact. The acronym if that “teachers are to DIE for”: that is they participate in *D*iagnosing the status of students as they begin lessons, they have multiple *I*nterventions that they can apply if their current intervention is not having the desired impact, and they *E*valuate the students’ responses to their interventions. It is less what teachers do in their teaching, but more how they think about their role. It is their mind frames, or ways of thinking about teaching and learning, that are most critical.

Table 1
Rankings and Effect-Sizes for 195 Influences on Student Achievement

Rank	Influence	ES	Rank	Influence	ES	Rank	Influence	ES
1	Teacher estimates of achievement	1.62	66	Small group learning	0.47	131	Divorced or remarriage	0.25
2	Collective teacher efficacy	1.57	67	Concentration/Engagement	0.45	132	Mainstreaming	0.24
3	Self-reported grades	1.33	68	Relative age within a class	0.45	133	Bullying	0.24
4	Piagetian programs	1.28	69	Professional development	0.45	134	Values/moral programs	0.24
5	Conceptual change programs	1.16	70	Computer assisted instruction	0.45	135	Illness	0.24
6	Response to intervention	1.07	71	Science	0.44	136	Religious schools	0.24
7	Teacher credibility	0.90	72	Early Intervention	0.44	137	Competitive vs. individual	0.24
8	Micro teaching	0.88	73	CAI with college students	0.44	138	Individualized instruction	0.23
9	Cognitive task analysis	0.87	74	Motivation	0.44	139	CAI in science	0.23
10	Classroom discussion	0.82	75	CAI with elementary students	0.44	140	Programmed instruction	0.23
11	Interventions for learning disable	0.77	76	Outdoor/Adventure programs	0.43	141	Summer school	0.23
12	Interventions for disabled	0.77	77	Teacher expectations	0.43	142	Finances	0.23
13	Teacher clarity	0.75	78	School size	0.43	143	Matching style of learning	0.23
14	Reciprocal teaching	0.74	79	Philosophy in schools	0.43	144	Exercise/relaxation	0.22
15	Feedback	0.73	80	Intelligent tutoring systems	0.43	145	Visual/audio-visual methods	0.22
16	Providing formative evaluation	0.68	81	Communication strategies	0.43	146	Teacher verbal ability	0.22
17	Acceleration	0.68	82	Exposure to Reading	0.42	147	Extra-curricula programs	0.21
18	Creativity programs	0.65	83	Comprehensive instructional	0.42	148	Class size	0.21
19	Self-questioning	0.64	84	CAI in writing	0.42	149	CAI in small groups	0.21
20	Concept mapping	0.64	85	Behavioral organizers	0.41	150	School cultural effects	0.20
21	Problem solving teaching	0.63	86	Goals	0.40	151	Aptitude/treatment interactions	0.19
22	Classroom behavioral	0.63	87	Social skills programs	0.40	152	Learning hierarchies	0.19
23	Prior achievement	0.63	88	After school programs	0.40	153	School counseling effects	0.19
24	Vocabulary programs	0.62	89	Cooperative learning	0.40	154	Co-/team teaching	0.19
25	Time on Task	0.62	90	Enrichment	0.39	155	Special college programs	0.18
26	Not labeling students	0.61	91	Career Interventions	0.38	156	Within class grouping	0.18
27	Spaced vs. Mass Practice	0.60	92	Psychotherapy programs	0.38	157	Family structure	0.18
28	Teaching strategies	0.60	93	Gaming/simulations	0.37	158	Web-based learning	0.18
29	Direct Instruction	0.60	94	Music-based programs	0.37	159	Personality	0.17
30	Repeated Reading programs	0.60	95	Drama/Arts Programs	0.37	160	Teacher immediacy	0.16
31	Study skills	0.60	96	Worked examples	0.37	161	Adopted children	0.16
32	Pre-term birth weight	0.59	97	Reducing anxiety	0.36	162	Home-school programs	0.16
33	Spelling programs	0.58	98	Student centered teaching	0.36	163	Out of school curricula	0.15
34	Tactile stimulation programs	0.58	99	Creativity	0.35	164	Sentence combining programs	0.15
35	Service learning	0.58	100	Attitude to content domains	0.35	165	Distance education	0.13
36	CAI with learning needs students	0.57	101	Inquiry-based teaching	0.35	166	Problem-based learning	0.12
37	Mastery learning	0.57	102	Bilingual Programs	0.35	167	Ability grouping	0.12

Table 1 (*continued*)

Rank	Influence	ES	Rank	Influence	ES	Rank	Influence	ES
38	Pre school with at risk students	0.56	103	Decreasing disruptive behavior	0.34	168	Diet	0.12
39	Visual-Perception programs	0.55	104	Various teaching on creativity	0.34	169	Juvenile delinquent programs	0.12
40	Peer tutoring	0.55	105	Adjunct aids	0.34	170	Teacher education	0.12
41	CAI in other subjects	0.55	106	Pre school programs	0.33	171	Diversity of students	0.11
42	Cooperative vs. individualistic	0.55	107	Head start programs	0.33	172	Mentoring	0.09
43	Interactive video methods	0.54	108	Principals/School leaders	0.33	173	Subject matter knowledge	0.09
44	Socioeconomic status	0.54	109	Inductive teaching	0.33	174	School calendars/timetables	0.09
45	Classroom cohesion	0.53	110	Ethnicity	0.32	175	Detracking	0.09
46	Meta-cognitive strategies	0.53	111	Online, digital tools	0.32	176	Perceptual-motor programs	0.08
47	Comprehension programs	0.53	112	Teacher effects	0.32	177	Single sex schools	0.08
48	Scaffolding	0.53	113	Drugs	0.32	178	Gender on achievement	0.08
49	Cooperative vs. competitive	0.53	114	Systems accountability	0.31	179	Charter schools	0.07
50	Peer influences	0.53	115	Ability grouping for gifted	0.30	180	Sleep	0.07
51	Frequent/Effects of testing	0.52	116	CAI in mathematics	0.30	181	Whole language	0.06
52	Phonics instruction	0.52	117	CAI with high school students	0.30	182	Types of testing	0.06
53	Classroom management	0.52	118	Collaborative learning	0.29	183	College halls of residence	0.05
54	Home environment	0.52	119	Mobile phones	0.29	184	Multi-grade/age classes	0.04
55	Teacher-Student relationships	0.52	120	Homework	0.29	185	Parental employment	0.03
56	Play Programs	0.50	121	Home visiting	0.29	186	CAI in distance education	0.01
57	Second/Third chance programs	0.50	122	Desegregation	0.28	187	Student control over learning	0.01
58	Parental involvement	0.49	123	Early intervention in the home	0.27	188	Open vs. Traditional	0.01
59	Mathematics	0.49	124	Teaching test taking	0.27	189	Summer vacation	−0.02
60	Writing programs	0.49	125	Use of calculators	0.27	190	Welfare policies	−0.12
61	Questioning	0.48	126	CAI in reading/literacy	0.26	191	Retention	−0.17
62	School effects	0.48	127	Volunteer tutors	0.26	192	Television	−0.18
63	Self-concept	0.47	128	Use of powerpoint	0.26	193	Home corporal punishment	−0.33
64	Integrated curricula programs	0.47	129	Teaching reforms	0.25	194	Mobility	−0.34
65	Student rating of teaching	0.47	130	Early intervention	0.25	195	Depression	−0.42

Methods of Teaching

A feature of Table 1 is that the various methods of teaching are spread throughout the distribution. Some are quite powerful, whereas others have limited effect. There are practices of teaching that do have systematically high effects, and these include conceptual change programs, teachers being proficient at cognitive task analysis, and ensuring there are structured and deliberate opportunities for classroom discussion, problem solving, reciprocal teaching, practicing the skills over time (spaced), and involving students in the teaching of each other.

Many of these methods can be accomplished within the traditional lecture—regardless of the size of the class. In the school sector, one of the most contentious issues is the minimal effects of reducing class sizes ($d = .21$). This continues to surprise many, and the claim is that surely the mind frames and influences with the highest probabilities would be more readily implemented when class sizes are smaller. The major reason why the effects of reducing classes (from 600 to 60 to 30 to 15) has small but positive effects is that teachers rarely change how they teach when moved into different class sizes (Hattie, 2007). The issue of class size also is topical at the university level, with the greatest complaint being directed at the "lecture." There have been no meta-analyses on this topic; the closest is the synthesis of the Keller Plan (mastery learning $d = .53$), which has been shown to be a powerful model based on the traditional lecture. The Keller Plan involves dividing the material into short units, are often self-paced, and students taking summative assessments before being permitted to move to the next block of material.

The lecture is cost-efficient and most universities use larger classes via lectures to subsidize the smaller tutorial size classes (both at under and postgraduate level). Over the past few decades the nature of the traditional lecture has changed considerably for many. There are now many alternative ways to complement and modify the traditional format—use of clickers, quality of tutorials, constructive alignment of assignments, guided notes, and so forth—and other creative ways of engagement via lectures. Compare also the power and spread of Ted Talks—which is a variant of the traditional lecture.

Much of the effects can depend also on the student preparation for the lecture. Flipped classrooms today are similar to the previous requests for students to "read the chapter" before coming to class. Indeed "lecture" from Latin means "to read"—when professors 'read' or clarified "readings." The move to the flipped classroom (where they watch a video, try a set of exercises, or read some prereading before coming to class), although often overstated (and devoid of a lot of evidence about its effectiveness), should have a higher probability of success—provided that the students are aware of what success means or learn what success means early in the series of lessons, when they have high trust that finding out what they do not know leads to major opportunities to learn, when teachers take the opportunity to talk *with* and not so much *at* the students, and when there is feedback about what the students does and does not know when they enter the lecture (see deWitt, 2014). Flipping is but a precursor to consolidating this prior exposure to the skills and understanding to be learned.

Similarly, what happens after the lecture can be important. For example, Henk and Stahl (1985) conducted a meta-analysis that showed that students taking notes from lectures improved their learning modestly ($d = .34$), but reviewing these same notes increased learning dramatically ($d = 1.56$). It is not the taking but the review that matters. As can be seen in Table 1, structuring opportunities for students to then learn from each other, to practice over time, to receive feedback to correct errors and misconceptions, and to evaluate their learning are most valuable.

One of the major criticisms of the current research base, and current classrooms, is that the majority of studies are based on surface learning outcomes. Although surface learning is a critical first step toward deeper learning, this may not be sufficient. Perhaps the lecture is more effective in transmitting subject matter content, and other than "speaking to students" is needed to then move to deeper outcomes. Lectures can be opportunities for transmitting information, prioritizing this information, expressing an argument, and communicating a teacher's enthusiasm (or not) for the subject. Lectures also pose minimum threat to students—note for example the research by Watkins and Biggs (1996, 2001) relating to the "paradox of the Asian learner" where many Asian students use lectures to absorb the surface understanding to then use to subsequently move to the deeper learning. In comparison, many Western learners consider memorization of the lecture notes as sufficient to pass the course (and often they are reinforced in this belief by the nature of the assessments).

Many have argued that an alternative to the traditional lecture is problem-based learning (PBL). There are many variants of PBL, but essentially "students are presented with a problem in order to activate their prior knowledge. This prior knowledge is then built upon further as the learners collaborate in small groups to construct a theory or proposed mental model to explain the problem in terms of its underlying causal structure" (Schmidt, Rotgans, & Yew, 2011, p. 793). Then there follows a process of refinement of the students mental models via continued study, discussion, and resolution of problems. The claim is that students' interest will be aroused by the nature of the problem, as PBL acts as a motivating method to entice the learner to seek more knowledge and understanding to thence resolve the problem. They then conclude their review: "in summary, PBL seems to have fairly strong effects on learning and achievement compared with conditions in which learning is not driven by the presentation of problems" (p. 801).

There have been nine meta-analyses investigating the effectiveness of PBL, and most have been at the university level—particularly within first-year medical courses. The overall message appears to be that PBL is not a very effective method (overall $d = .08$; Table 2). But this is not necessarily the correct interpretation as it is likely that this failure is more a function of introducing PBL before students have sufficient surface knowledge. Albanese and Mitchell (1993) hinted that increased years of exposure to medical education could increase the effect of PBL and a lack of experience means that novice PBL students may have more errors in their knowledge base, add irrelevant material to their explanations, and engage in backward reasoning (from the unknown to the givens) whereas experts engaged in forward reasoning (from the givens to the unknowns—see also Gilhooly, 1990; Saunders et al., 1990). Walker and Leary (2009) also noted that novice PBL students tended to engage in a more backward driven

Table 2
Summary of Meta-Analyses Relating to the Effects of Problem-Based Learning

Study	Year	No. of studies	No. of students	No. of effects	*d*	Problem-based learning
Albanese & Mitchell	1993	11	2,208	66	0.27	PBL in medicine
Vernon & Blake	1993	8		28	−0.18	PBL in college level
Dochy, Segers, Van den Bossche, & Gijbels	2003	43	21,365	35	0.12	PBL on knowledge and skills
Smith	2003	82	12,979	121	0.31	PBL in medicine
Newman	2004	12		12	−0.30	PBL in medicine
Gijbels, Dochy, Van den Bossche, & Segers	2005	40		49	0.32	PBL on assessment outcomes
Walker	2008	82		201	0.13	PBL across disciplines
Schmidt, van der Molen, Te Winkel, & Wijnen	2009	10		90	−0.18	Constructivist problem-based learning on medical knowledge
Leary, Walker, Shelton & Fitt	2013	94	34,689	213	0.24	PBL

reasoning, which results in more errors during problem solving and may persist even after the educational intervention is complete. PBL in later years when students have acquired major concepts and basic knowledge may be much more successful than PBL in the first year of being involved in a profession. This is a hypothesis which needs further research investigation. It may be that as knowledge increases the effects of PBL also decreases (see Gijbels, Dochy, Van den Bossche, & Segers, 2005); and it does seem that PBL can have a marked negative effect on acquiring knowledge ($d = -.79$; Dochy et al., 2003). Although their data are from high schools, Schwerdt and Wuppermann (2011) compared lecture style teaching with problem-based learning. They found a 10 percentage point shift from problem solving to lecture style presentation resulted in an increase in student achievement of about 1% of a standard deviation. Their conclusion was that

> students taught by teachers, who devote more effective teaching time to lecture style presentation rather than letting students solve problems on their own or with the teacher's guidance, learn more (in terms of competencies tested in the TIMSS student achievement test). (p. 15)

It is also likely that the nature and structure of the "problem" may be critical. For example, Perkins (2014) noted that some projects are overly structured. "Students are told exactly what to do in the spirit of ensuring success, but then they get no experience of searching out conceptual tools or homing in on particular cases of application" (p. 118).

Another common method often seen as an influential alternative within higher education is using online or Internet based solutions. There have been 17 meta-analyses which have compared various forms of delivery of university courses (often online or distance courses with traditional on campus learning; Table 3). The overall effect-size ($d = .12$) strongly points to the conclusion—method does not matter. Much more important are the ways teachers, irrespective of method of delivery, make their success criteria clear, the degree of challenge and feedback, and the quality of the interactions among students and between students and the teacher. These may (or may not) occur during online (e.g., MOOCs) and in-person teaching.

Attributes of Teaching

It is likely that the method of teaching is less critical than the attributes of the teaching within the methods. For example, among the top influences (see Table 1) are expectations, response to intervention, collective teacher efficacy, feedback, and seeking formative evaluation of the impact of teaching. Again, care is needed as some of these influences have much variability. For example, feedback is a common denominator of many of these teaching attributes and it has one of the greatest variances across the many effect-sizes. Asking whether comments or grades provide more effective feedback, increasing the amount of comments, deciding on different methods to give the feedback, and increasing the amount of testing are the wrong feedback questions. It is more critical to be concerned with whether and the nature of feedback received by the students than the amount of feedback given; teachers tend to be more concerned with providing "where to next" and "how am I going" feedback whereas students judge the quality of feedback when it assist them with "where to next"; feedback is most powerful when it is sought, received, and reacted to by the teacher.

Table 3
Summary of Meta-Analyses Relating to the Effects of Evaluation of Teachers

Study	Year	No. of studies	No. of effects	*d*	Evaluation
Cohen	1980	22	22	0.33	Feedback from student ratings
Cohen	1981	19	19	0.68	Student rating of teacher
Cohen	1981	41	68	0.48	Student rating of teacher
Dowell & Neal	1982	6	6	0.40	Student rating of teacher
Abrami, Leventhal, & Perry	1982	12	12	0.29	Expressiveness of teacher
Cohen	1986	47	74	0.44	Student rating of teacher
Clayson	2008	17	42	0.66	Student rating of teacher

This power of feedback, or formative evaluation back to the teacher, is high on the rankings of effects. But it is the reception, reaction, and review of teaching in light of these student evaluations of their learning that is important. Students are very good evaluators of the impact of teaching on their learning. There have been seven meta-analyses of more than 141 studies on this topic and the overall average effect (d = .47) shows a high relation between their ratings and teacher effectiveness (see Table 4). There are fewer studies relating student ratings and actual learning gains, but the limited evidence is also very positive about the importance of seeking student evaluations of their learning as part of the evidence about the teachers impact (see Marsh, 2007). The major concern is not the accuracy (it is high) but whether the teachers actually use the information to enhance their teaching. Further, there can be some important moderators to the overall message: for example, the effects were higher for student ratings of teaching skill and teachers' knowledge of the subject than with teachers' rapport with students and lowest for ratings about the difficulty of the course.

That there are major influences that can enhance student learning points to the importance of training university academics to become more effective teachers, although there are no specific meta-analyses on the effectiveness of courses to enhance university teaching. The VL research points to the importance of ensuring that the university lecturer has the right mindset about teaching at this level. The mindset is not that students come to the class to be taught, but that the teacher comes to the class to evaluate the impact of their teaching. Informing the students of the criteria of success near the beginning of the course, having a clear understanding and modifying teaching in light of the students prior achievement, using teaching methods that focus on moving students from surface to deeper understanding, ensuring the assessment are aligned with the success criteria and appropriate proportions of surface and deep learning, maximizing feedback about "where to next,' and ensuring that the course is sufficiently challenging to entice the learners into aiming for mastery of the content and understandings desired. All these attributes can be taught.

There also is a key role in this process for concentrating on the assessment for the course, as this is where the students discover what is truly valued (as noted above, sometimes despite what the teachers say). Brown and Hirschfeld (2008) asked teachers about whether the learning desired in their courses was more surface or deep, and the majority said deep; he asked the students of these same teachers and they said surface. The students primarily used the nature of assessments to make their decision. It is thus important to ensure that there is constructive alignment between the course claims, the assessment requirements, and the students understanding of both. Biggs (1996, 2003) argued that constructive alignment has two premises: students construct meaning from what they do to learn, and the teacher aligns the planned learning activities with the learning outcomes. Thus, any course needs to be designed so that the learning activities and assessment tasks are aligned with the learning outcomes that are intended in the course.

This implies that students should not be expected to "work out" what is to be learnt and what it means to be successful in that learning (often they only learn this when they get the results of assessment back), but instead it is necessary to make these criteria of success clear before any teaching or assessments. Without such alignment the powerful effects of feedback, reporting from assessment, and self-regulated learning are less likely to occur.

Concluding Comments

The synthesis of the 1200+ meta-analyses certainly point to the student as the greatest source of variance in learning. The estimates from the synthesis are that about 50% of the variance in learning is a function of what the student brings to the lecture room or classroom (see Figure 2). They differ greatly, they bring different attributes and prior knowledge, they have different motivations and purposes for learning, they study in varied ways, some are collaborators some are loners, they have a manifold of likes and dislikes, and they can be bright or struggling. The next, and greatest source of variance that we have some control over, is the qualities of the teacher. About 20% to 25% of the total learning variance is in the hands of teachers—and as every student knows, teachers can differ markedly. Understanding the essence of successful teachers and teaching has

Table 4
Summary of Meta-Analyses Relating to the Effects of Mode of Delivery

Study	Year	No. of studies	No. of students	No. of effects	*d*	Distance education
Machtmes & Asher	1987	19		19	−0.01	Effectiveness of telecourses
Cavanaugh	1999	19		19	0.13	Interactive distance learning on achievement
Cavanaugh	2001	19	929	19	0.15	Interactive distance education
Shachar & Neumann	2003	72	15,300	86	0.37	Distance vs. traditional teaching
Allen, Mabry, Mattrey, Bourhis, Titsworth, & Burrell	2004	25	71,731	39	0.10	Distance vs. traditional classes
Cavanaugh, Gillan, Kromrey, Hess, & Blomeyer	2004	14	7,561	116	−0.03	Distance in all classes
Williams	2004	25		34	0.15	Distance in allied health science programs
Bernard, Abrami, Lou, Wozney, Borokhovski, Wallet, Wade, & Fiset	2004	232	3,831,888	688	0.01	Distance education
Bernard, Lou, Abrami, Wozney, Borokhovski, Wallet, Wade, & Fiset	2004	155		155	−0.02	Presence or not: Asynchronous and synchronous
Allen, Bourhis, Mabry, Burrell, & Timmerman	2006	54	74,275	54	0.09	Distance vs. traditional teaching
Lou, Bernard, & Abrami	2006	103		218	0.02	Distance education in undergraduates
Zhao, Lei, Yan, Lai, & Tan	2008	51	11,477	98	0.10	Distance vs. traditional classes
Bernard et al.	2009	34	32,000	74	0.39	Most interactive distance vs. least interactive distance
Jahng, Krug, & Zhang	2011	20	1,617	20	0.02	On line vs. face to face
Roberts	2011	59		86	0.55	Web-based distance learning
Ungerleider & Burns	2003	8	1,324	8	0.00	Networked vs. traditional
Jahng, Krug, & Zhang	2007	20	1,617	20	0.02	Online distance ed

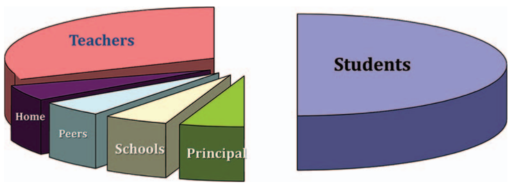

Figure 2. Distribution of variance of the various achievement influences. See the online article for the color version of this figure.

been the quest of the Visible Learning series. The remaining sources of variance relate to structural issues (lecture vs. online, nature of architecture of the classroom), peer effects (which could well be made to have greater impact that they do now), and leadership.

Since the first publications of VL, the aim has been to implement the major messages into schools and (more recently) universities. It became clear, very early, that the ways that teachers think about their role is the critical moderator. In some cases, the VL model is to give permission to those teachers who already have high impact on their students. In other cases, it entails understanding the beliefs about teaching and then work to modify them. Most university teachers come to the role with very strong beliefs about teaching (remembering that most were very successful in school and believe others can be also be successful if only they apply themselves, or if only they receive similar teaching to what they received). The VL research has identified eight major mind frames of successful teachers.

The most critical mind frame is "know thy impact"—when an academic walks into a teaching situation their fundamental question needs to be "how will I know my impact today." This leads to the three subquestions—what do I mean by impact today and have I communicated this to my students, what is the magnitude of the impact I am seeking, and how many students can I teach such that they attain this magnitude on the impact I have clearly communicated. The other critical mind frames follow from this first "Know thy impact." The others include the following: I am a change agent; I explicitly inform students what successful impact looks like from the outset; I see assessment as providing feedback about my impact; I work with other teachers to develop common conceptions of progress; I engage in dialogue not monologue; I strive for challenge and not 'doing your best'; I use the language of learning; and I see errors as opportunities for learning. The Visible Learning Into Action (Hattie, Masters, & Birch, 2015) outlines in more detail the implementation model to achieve the highest possible impact, and Clinton and Hattie (2014) explore in more detail the notion of teachers as evaluators.

This notion of educators as evaluators

> implies deliberate change, directing of learning, and visibly making a difference to the experiences and outcomes for the students (and for the teachers)—and the key mechanism for this activation is via a mind frame that embraces the role of evaluation. (Hattie & Clinton, 2011, p. 99)

It is very difficult, therefore, to see each teacher as an island implementing VL; instead it requires much intervention at the department and university level. University heads need to legitimize the debates about impact in their departments, create a trusting and fair environment for this to occur, and provide the resources (e.g., time) for collective collaborations in this pursuit of the highest possible impact on all students. As Senge (1990) noted, this requires leaders who are sincerely interested in the world of practice, who are highly respectful of that world, and are sincerely interested in using evaluative methods to make the education better for all. The leader's role also involves bringing expertise into evaluating the various interventions, developing partnerships, engaging communities, developing evaluation capacity, brokering knowledge into the university, forward thinking, and continuously attending to quality.

Faculty need to go beyond merely collecting data, creating reports, and asking students to fill in surveys, but to become excellent *interpreters* of evidence about their impact. This emphasis on interpreting the evidence about the teacher's impact is the major skill in being an effect teacher. Although outside consultants or commercially available sources can provide data reports, they will lead to little impact unless they are correctly interpreted by the teachers in a timely manner. As a consequence, "teachers as evaluators" adapt their teaching to maximize student learning. There needs to be a culture in departments and universities of seeking evidence to support interpretations about impact, having collective discussions about this impact, what the magnitude of this impact should be, and how pervasive is this impact on the students. The fundamental premise of Visible Learning is that when educators focus on defining, evaluating, and understanding their impact this leads to maximizing student learning and achievement. A key question is how to build the capacity of teachers and university administrators to collectively build and evaluate successful teaching programs and learning experiences (for a more detailed explanation of the evaluator mind frame, see Clinton & Hattie, 2014). There is a rich literature on how to be an effective teacher at the university level. It is obvious that we must develop and use the Sot and not let opinions (statements without evidence), fads, or favorite methods dominate the debates about what makes the difference to student learning.

References

Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine, 68,* 52–81. http://dx.doi.org/10.1097/00001888-199301000-00012

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32,* 347–364. http://dx.doi.org/10.1007/BF00138871

Biggs, J. (2003). *Aligning teaching for constructing learning*. York, UK: Higher Education Academy.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley. http://dx.doi.org/10.1002/9780470743386

Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice, 15,* 3–17. http://dx.doi.org/10.1080/09695940701876003

Chickering, A. W. (1969). *Education and identity*. San Francisco, CA: Jossey-Bass.

Clinton, J. M., & Hattie, J. A. C. (2014). Education and empowerment evaluation. In D. Fetterman, S. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self-assessment, evaluation capacity building, and accountability*. Thousand Oaks, CA: Sage.

DeWitt, P. M. (2014). *Flipping leadership doesn't mean reinventing the wheel*. Thousand Oaks, CA: Corwin Press.

Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction, 13,* 533–568. http://dx.doi.org/10.1016/S0959-4752(02)00025-7

Gijbels, D., Dochy, F., Van den Bossche, P., & Segers, M. (2005). Effects of problem-based learning: A meta-analysis from the angle of assessment. *Review of Educational Research, 75,* 27–61. http://dx.doi.org/10.3102/00346543075001027

Gilhooly, K. J. (1990). Cognitive psychology and medical diagnosis. *Applied Cognitive Psychology, 4,* 261–272. http://dx.doi.org/10.1002/acp.2350040404

Hattie, J. A. C. (2007). The paradox of reducing class size and improved learning outcomes. *International Journal of Education, 42,* 387–425.

Hattie, J. A. C. (2009). *Visible learning: A synthesis of 800 meta-analyses relating to achievement*. Oxon, England: Routledge.

Hattie, J. A. C. (2011). Which strategies best enhance teaching and learning in higher education? In D. Mashek & E. Hammer (Eds.), *Empirical research in teaching and learning: Contributions from social psychology* (pp. 130–142). Hoboken, NJ: Wiley-Blackwell. http://dx.doi.org/10.1002/9781444395341.ch8

Hattie, J. A. C. (2012). *Visible learning for teachers. Maximizing impact on achievement*. Oxford, UK: Routledge.

Hattie, J. A. C., & Anderman, E. (2013). *Handbook on student achievement*. New York, NY: Routledge.

Hattie, J. A. C., & Clinton, J. M. (2011). School leaders as evaluators. In *Activate: A leader's guide to people, practices and processes* (pp. 93–118). Englewood, CO: The Leadership and Learning Center.

Hattie, J. A. C., Masters, D., & Birch, K. (2015). *Visible learning in action*. Oxford, UK: Routledge.

Hattie, J. A. C., & Yates, G. (2014). *Visible learning and the science of how we learn*. London, UK: Routledge.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.

Henk, W. A., & Stahl, N. A. (1985, November). *A meta-analysis of the effect of notetaking on learning from lecture. College reading and learning assistance*. Paper presented at the Annual Meeting of the National Reading Conference, St. Petersburg Beach, FL.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 319–383). Amsterdam, The Netherlands: Springer. http://dx.doi.org/10.1007/1-4020-5742-3_9

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: Vol. 2. A third decade of research*. San Francisco, CA: Jossey-Bass.

Perkins, D. (2014). *Future wise: Educating our children for a changing world*. Hoboken, NJ: Jossey-Bass.

Saunders, N. A., Mcintosh, J., Mcpherson, J., & Engel, C. E. (1990). A comparison between University of Newcastle and University of Sydney final-year students: Knowledge and competence. In Z. H. Nooman, H. G. Schmidt, & E. S. Ezzat (Eds.), *Innovation in medical education: An evaluation of its present status* (pp. 50–54). New York, NY: Springer.

Schmidt, H. G., Rotgans, J. I., & Yew, E. H. (2011). The process of problem-based learning: What works and why. *Medical Education, 45,* 792–806. http://dx.doi.org/10.1111/j.1365-2923.2011.04035.x

Schwerdt, G., & Wuppermann, A. C. (2011). Is traditional teaching really all that bad? A within-student between-subject approach. *Economics of Education Review, 30,* 365–379. http://dx.doi.org/10.1016/j.econedurev.2010.11.005

Senge, P. (1990). The leaders new work: Building learning organizations. *Sloan Management Review, 32,* 7–23.

Tomcho, T. J., & Foels, R. (2008). Assessing effective teaching of psychology: A meta-analytic integration of learning outcomes. *Teaching of Psychology, 35,* 286–296. http://dx.doi.org/10.1080/00986280802374575

Walker, A., & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-based Learning, 3,* 12–43. http://dx.doi.org/10.7771/1541-5015.1061

Watkins, D. A., & Biggs, J. B. (2001). The paradox of the Chinese learner and beyond. In D. A. Watkins & J. B. Biggs (Eds.), *Teaching the Chinese learner: Psychological and pedagogical perspectives* (pp. 3–26). Hong Kong, CERC: ACER.

Watkins, D. A., & Biggs, J. B. (Eds.). (1996). *The Chinese Learner: Cultural, Contextual and Psychological Influences*. Hong Kong, CERC: ACER.

Received January 3, 2015
Revision received January 21, 2015
Accepted January 21, 2015 ■