Article

The Long View of *Visible Learning*'s Impact

Douglas Fisher * and Nancy Frey

Department of Educational Leadership, San Diego State University, San Diego, CA 92182, USA; nfrey@mail.sdsu.edu

* Correspondence: dfisher@mail.sdsu.edu

check for updates

Abstract: In this article, we address the common criticisms of the Visible Learning research and offer a long-term view of the potential presented with this body of knowledge. We contextualize our view with some experiences in a high school that is focused on improving student learning.

1. Introduction

When the book *Visible Learning* (Hattie, 2009) [1] was published, *The Times*' Educational Supplement suggested that this compilation of educational research "reveals teachings' Holy Grail". We have not interviewed anyone from *The Times*, but we suspect that they identified the body of work as the "Holy Grail" because it represented the largest collection of educational evidence ever assembled. At the time, the database included over 800 meta-analyses that included 50,000 research studies involving more than 150 million students. The database has grown since it was first published to include over 1600 meta-analyses consisting of 95,000 research studies involving more than 300 million students. To say that this is impressive would be an understatement. As such, we recognize that there has been both celebrations and criticisms about this body of work. In this article, we review and acknowledge the critics' perspectives (e.g., Myburgh, 2016 [2]) and then offer our take on the enduring messages gleaned by adopting a long-view of the evidence summarized and synthesized in *Visible Learning*. As practitioners who have worked to implement the findings from this body of work, we offer reflections about the lessons we have learned.

2. Critics of Visible Learning

The claim that *Visible Learning* revealed the Holy Grail invited critics to the table. In this article, we focus on some of the more substantial and thoughtful questions posed about the information contained in *Visible Learning*. We are not the authors of this body of work nor the keepers of it, but rather teacher-scholars who have worked to understand the implications of the research in an effort to help schools improve. We want to acknowledge that we believe that *Visible Learning* has been a valuable resource when used correctly.

A common concern about the *Visible Learning* (VL) evidence base focuses on the use of meta-analyses themselves. There is concern that the very tool that grounds *Visible Learning* is flawed. The risk with the tool is that big facts are identified but context and complexity are lost. As Glass (2000) [3] (p. 9) commented, "averages do not do it justice". Philippa Cordingley and colleagues describe meta-analyses as "[t]he outcome is a kind of high-level map of the evidence—good enough to find the island but not accurate enough to lead you to the buried treasure" (http://www.curee.co.uk/node/5109) [4]. One of the reasons for inventing meta-analysis was to ask about the various factors that may moderate the average effects—and contrary to many critics' claims, these are elaborated extensively throughout VL.

To our thinking, that's the whole point of summarizing large bodies of work: to identify the big ideas. And then to note when they do, and do not, work. Meta-analyses are a tool, like many other statistical tools, to summarize data. It's not the only way, but it does allow for effect sizes to be calculated, thus reducing the impact of sample size on statistical significance for individual studies. We have been in countless numbers of conversations in which one study is compared with another and the findings and recommendations differ. As a result, the adage that "research says" becomes, "there is a study that proves everything" and no progress is made in recognizing what works best in improving learning outcomes for students.

As Hattie predicted, some of the criticism focuses on the ways in which studies are combined. As he noted (2009) [1], there would be concerns about combining "apples and oranges" (p. 10). In some cases, the outcome measures differ even though the focus of the study was similar. For example, in studies of writing improvement, some researchers use state writing scores, others use rubrics that have been validated, and still others use analytic scores. The concern is that one is apples, another is oranges, and another is bananas and that they can't be combined. But the fact is that they are studying writing, albeit with different measures, and we should be able to say something about writing improvement. As Hattie notes, apples and oranges are still fruit, and if we can accept that we're studying fruit, then the tool works. *Visible Learning* does not combine studies of writing with studies of class size. That would be like combining fruit and furniture.

Another criticism we have heard concerns the studies that comprise the meta-analyses. In these cases, it seems that the concern is about the ability of individual studies to influence the overall message, or as some have said *quality in, quality* out versus *garbage in, garbage out*. To a large extent, this is true. There is a range of quality in all human endeavors, from the hair stylist to the mechanic to the teacher to the researcher. And it could be that individual researchers included slightly lower quality, or much higher quality, studies in their meta-analysis. Hattie did say there were many other sources to read about the impact of quality on the meta-analyses so *Visible Learning* was not going to dwell on this matter, but still he made calls about the quality of some of the meta-analyses. In addition, the findings are averaged over large numbers of studies, thus minimizing the impact of any one study. And we have to believe that people are making a good faith effort to contribute to the knowledge base. When the numbers get to 300 million students, it's hard to believe that a poor quality study or two will sway the results in a significant way.

Along those same lines, there have been concerns that some individual studies are included in more than one meta-analysis, thus increasing the impact of some studies relative to others. This is the nature of the tool being used, which is why the database has to be so large as to identify trends. It's also why we focus on influences that are above average rather than saying x is better than y because it has an effect size that is 0.06 higher. If it were a perfect world, each study would only be included in one meta-analysis, but it's not and we have to accept the evidence as it is presented.

In addition, some criticize the fact that the effect sizes change over time. For example, teacher-student relationships used to be 0.72 and more recently they are listed at 0.52. Critics have used the changing effect sizes to suggest that the database is not stable and that the influences are variable. It is important to understand why there can be such changes. In this case, for example, it is because a more recent study (Vandenbroucke, Spilt, Verschueren, Piccinin, & Baeyens, 2018 [5]) asked about the effects of teacher–student relations on executive functioning, working memory, inhibition, and cognitive flexibility, which are among the critical achievement outcomes (and they are low, $d = 0.18$). Unlike referencing a single published article that will maintain the effect size over time, the *Visible Learning* database is continually updated and the details of new articles can be critical. As new meta-analyses are published, they are integrated into existing evidence, which may result in changes in effect size. Over the past decade, there have been very few major changes to the story. In essence, the story remains the same. There are things that seem to work to accelerate student learning and there are things that are not very useful and perhaps even harmful.

Finally, there is concern that *Visible Learning* focuses on academic outcomes to the exclusion of other valuable impacts on students, such as social and emotional development or physical growth. Hattie (2009) [1] was clear that achievement is but one, very important, outcome of schooling. That is the lens of this database, which does not prevent other researchers from engaging in the same process to identify what works best in other pursuits. We take it at face value that schools exist, in large part, to ensure that students learn. Thus, we focus on students' academic progress.

3. Unintended Consequences from *Visible Learning*

As the saying goes, no good deed goes unpunished. And there have been some unintended consequences that result from the summary of the world's largest research compilation. One of those is the David Letterman effect. In the US, David Letterman hosted a late night TV show that included a regular segment that used humor and named the top 10 items on a list, such as "Top 10 Things Stupid Americans Say to Brits" or "Top 10 President Obama Excuses". In some school systems, the *Visible Learning* database has been reduced to a top 10 list and administrators focus on the highest effect size influences, irrespective of their complexity in implementation or value to the school. That's really not the point of the list. The point was to identify influences that were much more likely to accelerate students' learning, such that all students gain at least a year of learning for a year of schooling.

A second unintended consequence involves focusing on the short term. Far too many schools and districts have approached *Visible Learning* as the flavor of the day and provided teachers a one-shot overview of the evidence base. These sessions are replaced the next month with a different area of focus. As a result, teachers tend to focus on individual strategies (often ones that they are familiar with) rather than focusing on the learning that students are, or are not, doing.

4. The Long View

When schools and school systems move beyond discussions of individual influences that are included on the *Visible Learning* list and instead use the evidence to think longer-term about the changes that must be made to ensure that all students learn, several key ideas emerge. In this section, we present each of our long-view lessons with some anecdotal evidence from the school where we work, Health Sciences High in San Diego, CA, USA. We are not suggesting this is an empirical investigation or even a case study, but rather we use authentic examples from our school to highlight the ways in which teachers have interpreted the *Visible Learning* philosophy over several years. We interviewed 26 teachers about the changes they have experienced on our *Visible Learning* journey and use quotes from those interviews to explore each long view lesson learned.

4.1. Focus on Learning, Not Teaching

One of the major messages contained within *Visible Learning* centers on the idea that discussions at school should focus on students' learning rather than the instructional routines and procedures designed to increase learning. The tension, of course, is that the database includes a number of instructional routines, such as Reciprocal Teaching (ES = 0.74) and concept mapping (ES = 0.64). Too often, discussions among educators, or between teachers and their students, centers on the use of the strategy rather than the learning that results from the use of the strategy. Of course, teachers have to plan meaningful lessons that engage students in a variety of tasks, but the focus should remain on whether, or not, students learned anything from the experience.

At Health Sciences High, conversations among teachers used to center around the use of specific teaching strategies. For example, just a few years ago it was not uncommon for an instructional coach to focus discussions with teachers around instructional strategies they should be using. We remember one meeting in which the coach was guiding the conversation with a group of history teachers about a specific form of note-taking (ES = 0.50). The coach was encouraging everyone to use this type of note-taking, suggesting that students need to learn to get things written down so that they can study. The focus of the conversation was on the teaching, limited to one approach that seemed to have evidence for success.

Conversations today are very different. The discussions begin with teachers talking about their students' learning. For example, a group of science teachers started their discussion saying, "At least 86% of our students successfully included evidence for their claims. There were three main claims, each of which could be supported with evidence. We have a few students who made claims that could not be supported by the evidence and two students who made claims but used the wrong evidence. Overall, the learning levels where high. Remember when we started this unit, only about 10% of the students were able to write with claims, evidence, and reasons."

The conversations eventually move to the tools that were used to facilitate that learning as teachers identify gaps in students' accomplishments. During the science teacher conversation, one of the teachers noted, "I used a lot of peer tutoring [ES = 0.53] to focus my students. When I look at the difference in their writing over the past 10 weeks, it seems to have worked. What else do you all do to get the success levels up?"

In response, another teacher commented, "I used practice testing [ES = 0.54]. I gave students practice versions of the constructed response task and had them analyze their performance and then form study groups. My percentages of success are about the same as yours, so I guess these two approaches worked about the same. I wonder what would happen if we used both peer tutoring and practice testing? Would it result in even higher learning, or is there a ceiling to student performance based on the tools we are using?"

The conversation continued with teachers discussing students' learning and sharing the tools they used to get there. It may seem like a subtle difference, but focusing on learning first requires that teachers examine evidence of achievement. It also reinforces the idea that there is not one right way to teach, but there probably are less effective ways. It also ensures that teachers come to understand that they should not hold any instructional routine in higher esteem than their students' learning. Instead, the focus should remain on changing their practices if students fail to learn at acceptable levels.

This group of science teachers also talked about responding to the students who did not learn at high levels. As one of them commented, "I have about five students who still don't get it. I'm looking for advice about how to re-teach this. I want to make sure that they all get it before we move on. Do any of you have ideas? Or do you have plans for the students in your classes who didn't reach mastery?" The power in this part of the discussion is the ownership teachers experience for students' learning. Rather than thinking, *I taught it, but they didn't get it*, this teacher is saying, *I must ensure that all students learn at high levels and I need guidance from my peers to accomplish my goal*. If the conversations were limited to teaching strategies, we're not sure that this teacher, or any of his colleagues, would exhibit such ownership.

4.2. Know Thy Impact

Another long view from *Visible Learning* is the focus on impact. This saying from *Visible Learning* was not clearly understood a few years ago. The science teachers described above clearly value the impact they have on students learning. They see themselves as influencers of learning and note the impact they have had. But it wasn't always so. In the past, teachers at our school did not know how to determine their impact or assumed they had an appropriate level of impact based on the summative assessments they administered. It seemed as though they assumed they had an impact if students did well at the end of the unit, irrespective of how well they might have done at the outset of the unit. As one of our math colleagues now says, "I always assumed they learned it from me when I assessed their knowledge. But I now realize they could have known it before me and that I might waste a lot of time teaching things that my students already knew."

To determine impact, teachers have to understand students' current levels of understanding and performance. Then they have to measure change over time to assess the value of the learning experiences students have had. Both pre-assessment and post-assessment are required if teachers are going to claim impact. In addition, teachers need tools to calculate their impact on learning. Effect sizes are one tool that teachers can use to estimate their impact. Of course, we cannot use

Hattie's 0.40 average when using teacher-created assessments as the average is based on published meta-analyses, often using tools with strong psychometric properties. However, our teachers do use the Cohen's *d* method to determine their impact and note when the effect size is low versus high. It is fodder for conversation and action; it's comparing the effects across students and classes within Health Science High. It's not evidence for a research article that we're after.

In addition, the *Visible Learning* philosophy requires that teachers discuss amount of impact they expect to have. Teacher expectations, with an effect size of 0.43, are an important consideration. If teachers have low expectations, students will probably meet them. When teachers have higher expectations, combined with strong relationships with students and quality instructional experiences, students will likely achieve more. It was not common at our school (and probably is still not at many schools) for teachers to discuss the expectations they have for students. Presumably, the content standards establish the level of expectation for learning, but teachers have to unpack those expectations and internalize the level of rigor necessary for students to achieve well.

It is currently more commonplace for teachers to discuss their expectations for learning with their peers. For example, during a department meeting, a math teacher asked, "How much growth are we expecting from students in terms of linear functions?" As a result, the team analyzed the standard identifying the skills and concepts necessary for students to demonstrate mastery. When asked about this later, the teacher commented, "We never used to dig deeply into the standards to understand the expectations and necessary knowledge. Instead, we all assumed we all understood the standard and that the textbook would provide appropriate practice for students."

Thus, the concept of determining impact has changed the nature of the conversation such that teachers collect evidence at the outset of a unit of study, identity the appropriate impact necessary to ensure learning, and then determine the impact that they have had. Much like our first long view lesson, focusing on impact allows teachers to engage in conversations about the next steps necessary for students who do not initially reach competence. They seem to take more responsibility for students learning and leave less to chance. As one of the English teachers said, "It's humbling but also exciting. We all got into this profession because we wanted to impact the lives of young people. Now we're figuring out if we really did and, if not, what we can do about it before it's too late and they leave us. This really does ensure that we leave no child behind."

4.3. The value of Clarity

Spending time focused on expectations opened the door to conversations about teacher clarity (ES = 0.75). Fendick (1990) [6] defined teacher clarity as "a measure of the clarity of communication between teachers and students in *both directions*" (p. 10) and suggested that there were several aspects of clarity, including:

- Clarity of organization;
- Clarity of explanation;
- Clarity of examples and guided practice;
- Clarity of assessment of student learning.

At Health Sciences High, we focused more narrowly than Fendick, although all four of his points have been emphasized over the years. Our starting point was clarity of the learning intentions and success criteria. In teams, teachers identified appropriate learning intentions and success criteria for their lessons and then worked to communicate these expectations to their students. We defined the two terms as follows:

- **Learning Intentions** are statements that describe the intended learning outcomes of a lesson(s). They identify what the learners are expected to know, understand and be able to do as a result of the learning.
- **Success Criteria** specify what learners will do to demonstrate that they have met the Learning Intentions. They show the learner and teacher that learning intentions have been achieved.

Prior to our *Visible Learning* journey, teachers planned lessons based on objectives. Sometimes the objectives were written on the dry-erase board and sometimes they were not. They certainly did not drive teaching and learning. Instead, they were posted in compliance with misguided, but well-meaning, directives from administration. Teachers and students did not routinely discuss the objectives even though they were purportedly used to design learning experiences.

More recently, learning intentions and success criteria have replaced objectives on the walls in our school. Not in compliance to any directive, but rather based on an understanding that students deserve to know what they are expected to learn and how they will know if they learned it. In some cases, the learning intentions and success criteria are shared at the outset of the lesson. At other times, they are revealed during the lesson. One of the science teachers noted, "When we're engaged in inquiry, I don't want to ruin the experience by sharing the learning intention first. But I do think it's important that they know what they were supposed to learn from the experience some time during the lesson."

The learning intentions and success criteria are reviewed several times during a lesson and students are expected to reflect on the success criteria to determine whether or not they believe that they have reached the appropriate level of mastery. For example, an English teacher has students respond to a prompt in writing before they leave class each day. The prompt requires that students reflect on the success criterion/a for the day. On their way out the door, students place their "exit slip" in one of four boxes, labeled as follows:

- I am just learning and I need more help.
- I'm almost there and I need more practice.
- I own it and can work independently.
- I'm a pro and I can help others.

This process is one of many that teachers at Health Sciences High use to ensure that learning intentions and success criteria are visible, transparent, and obvious for students. Over the years, the focus on teacher clarity has allowed the faculty to hone their lessons, ensuring that the activities and tasks they plan for students provide them opportunities to practice and apply before being assessed. As one of the math teachers commented,

> These really do all fit together. Greater clarity requires an understanding of expectations. Expectations have to be communicated so that students accept the challenge of learning. And you have to determine the impact to make sure that you're focused on learning. And then plan experiences, choosing from a range of potentially appropriate strategies to have that impact. And then you have to respond when the impact isn't where you want it. It's a different approach than simply saying I'm going to teach this and I hope you learn it.

4.4. Learning Occurs in Phases (and Some Approaches Are More Effective Than Others at Each Phase)

Visible Learning introduced to some, and reintroduced to others, the idea that learning occurs across phases. The journey begins at the surface level and through the deep level, eventually (hopefully) resulting in transfer. It's a process from surface to deep to transfer of learning. Actually, it's probably better to say that it's a phase because learning starts at the surface level and then, with the right experiences, can move to the deep level. And then with different experiences can result in transfer. As with other long view lessons learned from *Visible Learning*, this was not a common conversation at Health Sciences High. Rather, teachers tended to focus on today's lesson and its alignment with standards, not considering the phases a learner must traverse to own the learning. A history teacher noted,

> I hadn't ever heard of the phases of learning. I just planned tasks that seemed to be aligned with the standards. I know that sometimes my students struggled too much so I had to tell them things so that they could complete the task, but I didn't consider that it was because they needed more surface learning before they could engage in deep learning.

One of the math teachers commented,

I was worried when I first heard about surface learning. I thought it was superficial and procedural. I want my students to be deeper thinkers and I didn't think I should spend time on surface learning. I also thought that surface learning would involve only direct instruction and teacher telling. Boy was I wrong. Students need to have surface learning if they are going to go deep. I want my students to be able to complete rich tasks, but these tasks require foundational knowledge and that is more the procedures and memorization. They have to know stuff and be able to use that stuff. For deep learning, students have to identify relationships between ideas, see the connections, and develop their schema. The number of minutes I spend on surface learning varies, based on my students' needs. The difference now is that all of our units of study acknowledge the various phases of learning. We also assess students at each of the phases and use that information to determine if our instruction has had the desired impact.

In addition to recognizing the phases of learning and designing lessons and assessments to match the phases, teachers have studied the various instructional tools that are likely to be useful at each phase of learning. The risk here is that we focus too much on teaching and not learning, which has been a regular part of the conversations about surface, deep, and transfer learning.

A group of eight teachers worked together to define each phase, identify a driving question, name the processes in that phase, and then suggest sample instructional routines that might develop students' expertise at that phase. The results of their efforts can be found in Table 1. This tool helps teachers plan their instruction. But even more importantly, it sends a strong message that progress along the phases of learning are influenced by the teacher and the actions the teacher takes. It also allows teachers to monitor students' progress and ensure that learning from these experiences remains the goal.

4.5. Student Performance Is Feedback to Us

The final long view lesson learned from years of following the Visible Learning journey was also clearly outlined in the 2009 book. As Hattie noted, feedback is effective. But it's more than giving feedback, especially corrective feedback, to students. Instead, student performance should be seen as feedback to the teachers, allowing them to reflect on their lessons, determine the impact that they have had, and design future lessons that address more needs.

This has been a hard lesson to learn as all of us, at some point, have focused on student factors that interfere with their learning. Faculty meeting time used to stray to students' home lives, their poverty, the language learning needs, their motivation, their attendance, and a host of other things. We're not saying that these are unimportant factors, but rather they detract from the idea that student performance should be seen as feedback to the teacher about the lesson and the next steps for learning.

As an English teacher said,

Honestly, if we really think about it, I think we used to blame the victim a bit or engage in *pobrecito* syndrome [poor baby], feeling sorry for students because of their life circumstances. We can be concerned, and notice when they need help from a counselor, and still focus on their learning. When we learn to see student work as a feedback to us, the world changes. When my students do not do well on an essay, I don't blame them. I don't say they were lazy or that their parents don't care about school. I reflect on what I could have done and what I will do to get the outcomes that my students deserve. It's not that I beat myself up, or that someone thinks that I'm not effective, but rather that I understand that I have a profound role in students learning and I choose to take responsibility for that. I want to have an impact on students and their performance provides me with data, information, and feedback that I need so that I can get better.

Table 1. Phases of Learning and Associated Instructional Routines.

Learning Phase	Definition	Driving Question	Processes	Content Literacy Routines and Effect Size (*d*) from Hattie
Surface Learning	Acquisition and consolidation of initial knowledge base	What are the key facts and principles?	Rehearsal, memorization, and repetition	• Leveraging prior knowledge (*d* = 0.67) • Vocabulary techniques (sorts, word cards, mnemonics, etc.) (*d* = 0.67) • Reading comprehension in context (*d* = 0.60) • Wide reading on the topic under study (*d* = 0.42) • Summarizing (*d* = 0.59)
Deep Learning	Interaction with skills and concepts	How do these facts and principles fit together?	Planning, organization, elaboration, and reflection	• Concept mapping (*d* = 0.60) • Discussion and questioning (*d* = 0.82) • Reciprocal teaching (*d* = 0.74) • Metacognitive strategies (*d* = 0.69)
Transfer Learning	Organizing, synthesizing, and extending conceptual knowledge	How and when do I use this for my own purposes?	Making associations across knowledge bases Application to novel situations	• Reading across • Documents to conceptually organize (*d* = 0.85) • Problem-solving teaching (*d* = 0.61) • Peer tutoring (*d* = 0.55) • Formal discussion (debate, Socratic seminar) (*d* = 0.82) • Extended writing • (*d* = 0.43)

5. Conclusions

There is a reason that a book published more than a decade ago is still being discussed in educational circles. It has something to say that matters. Although there are critics, and it's important to hear those voices, the main messages are germane today. It's not simply that teachers matter, but rather how teachers think that matters. There are some things in our profession, such as grade-level retention (ES = −0.32] that are harming students. There are other things that really don't matter much, such as class size reduction (ES = 0.21), even though they are taking up space in conversations about school improvement. In addition, there are things that really seem to accelerate learning, such as response to intervention (ES = 1.29) and developing collective teacher efficacy, that probably need to be in place in every school.

We think there is more to *Visible Learning* than the isolated effect sizes calculated from the meta-analyses. To our thinking, the enduring messages contained within the *Visible Learning* story are not about one specific instructional routine being more useful than another. Instead, the long-view lessons are about focusing on learning, understanding one's impact, ensuring clarity, understanding the science of learning, and using student performance as feedback for teachers. When these become commonplace philosophical stances in schools, we believe students will be more successful. When teachers develop the habits required to enact these five stances, their impact will grow and it's anyone's guess on what the students can achieve. *Visible Learning* is more than a dataset; it's a way of thinking about the work we do.

References

1. Hattie, J. *Visible Learning: A Synthesis of over 800 Meta-Analyses Relating to Achievement*; Routledge: New York, NY, USA, 2009.
2. Myburgh, S.J. *Critique of Peer-Reviewed Articles on John Hattie's Use of Meta-Analysis in Education*; Department of Education Working Papers Series; University of Bath: Bath, UK, 2016.
3. Glass, G.V. Meta-Analysis at 25. 2000. Available online: http://glass.ed.asu.edu/gene/papers/meta25.html (accessed on 18 October 2018).
4. Systematic Reviews and Weather Forecasts. Available online: http://www.curee.co.uk/node/5109 (accessed on 18 October 2018).
5. Vandenbroucke, L.; Spilt, J.; Verschueren, K.; Piccinin, C.; Baeyens, D. The classroom as a developmental context for cognitive development: A meta-analysis on the importance of teacher-student interactions for children's executive functions. *Rev. Educ. Res.* **2018**, *88*, 125–164. [CrossRef]
6. Fendick, F. The Correlation between Teacher Clarity of Communication and Student Achievement Gain: A Meta-Analysis. Unpublished Doctoral Dissertation, University of Florida, Gainesville, FL, USA, 1990.