

Debunking Hattie: Evaluating the Contribution of Academic Studies to Policy Development and Implementation in Australia

Professor Eddie Blass

Torrens University

Australia

&

CEO

Inventorium Pty Ltd

Abstract

John Hattie's book 'Visible Learning' published in 2009 can only be described as an amazing feat of scientific inquiry. He and his team created a mega-analysis, a meta-analysis of 800 meta-analyses, to see what the visible effect sizes were of different elements on education. To generalise, pretty much everything had an effect, mainly positive, with the mean effect size being 0.4. This has then been pretty much adopted into practice as anything being above 0.4 as having an above average impact, and hence being 'good', and anything below 0.4 having a below average impact and hence being 'not so good'. While such a sweeping generalisation 'makes sense' in terms of rational logic, I argue here that it has been incredibly damaging in limiting the thinking on educational progress and change in the Australian policy context. I stage this argument in 5 sections: methodology, context, assumptions, leadership and currency – and offer this as a framework for evaluating academia's contribution to policy agendas. I conclude by reflecting on Hattie's later papers which move his conclusions forwards academically but have yet to do so in practice.

Keywords: education policy; evaluation; meta-analysis

Introduction

John Hattie's book 'Visible Learning' published in 2009 can only be described as an amazing feat of scientific inquiry. He and his team took 800 meta-analyses carried out in education, and created a mega-analysis, a meta-analysis of the meta-analysis, to see what the visible effect sizes were of different elements on education. To generalise, pretty much everything had an effect, mainly positive, with the mean effect size being 0.4. This has then been pretty much adopted into practice as anything being above 0.4 as having an above average impact, and hence being 'good', and anything below 0.4 having a below average impact and hence being 'not so good'.

What is clear from Hattie's analysis is that the teachers are the most influential factor on a child's performance, and the notion that teachers must be able to take the perspective of their students, and students need to see themselves as their own teachers is a core premise to the idea of visible learning. However, the statistical mega-analysis suggests that the pedagogic approaches that best support this stance are the ones which show weakest effect on improving performance, eg inquiry based learning and project based learning, and Hattie himself berates the teacher as a facilitator model, preferring instead the more didactic teacher as instructor, constantly testing kids' performance.

This view has largely been accepted without question, despite the fact it is somewhat paradoxical. Instructor led learning and students seeing themselves as their own teachers could be contradictory, unless the instruction itself is in the learning process which returns us to PBL and inquiry-led learning – the very pedagogical approaches Hattie criticizes. Rather than exploring this peculiar juxtaposition to make sense of it, Hattie's views have been accepted largely without question across Australian schools and this has limited pedagogical innovation severely. Indeed, Australian teacher training does not reward innovation in the classroom, but rather regulation and standardisation; and the Teacher Registration Boards require hours of continuous professional development rather than innovative development. Imagine, for example, if each teacher had to evidence 5 hours of innovative practice per year where education might be now in the classroom!

This paper seeks to 'debunk' the impact that Hattie's work has had on stifling innovation in Australia. If you read the whole book – it is clearly not what was intended; and the use of elements of his work in isolation is akin to baking a cake with half the ingredients. The cake will not rise nor be edible, but that is what has happened in Australia as the response to this great piece of work.

I am not trying to undermine or devalue the work of Hattie and his team – it is an incredible feat, and a very interesting academic exercise. But as with all things academic, in order for them to have meaning in real life, one needs to consider the context, the assumptions underpinning the conclusions, leadership of implementation, and currency of the work (which relates back to context). Simply having a recipe is not sufficient.

You need the kitchen (context), the ingredients (assumptions), the chef (teachers and leadership) and the currency (ingredients within use by date) in order to actually bake the perfect cake. And even then, the proof of the recipe is in the eating.

The first element to be discussed is the methodology – and this is not, as I reiterate, to undermine it, but is more to understand how the methodology impacts on the findings, and how this in turn relates to the four following elements.

Methodology

Terhart (2011:434) applauds Hattie's book for the 'very broad perspective, his resolute manner of bundling results which have already been bundled, his calculation of (cumulative) effects (although not available in detail), his visually appealing way of presenting each effect size by 'barometers', his ability to synthesize broad quantitative information in a verbally convincing manner and to draw inspiring theoretical conclusions from the material—all make the book a veritable treasure for researchers since it presents and comments important information for those theoretically and practically interested in teaching, teachers, classrooms, and schools.'

However, he questions the methodology of the mega-analysis (ibid: 427) 'Depending on data quality and data analysis, effect size can be calculated in very different ways. Concerning this problem and also concerning his mode of aggregating already aggregated effect size calculations (!) Hattie gives no detailed information.' In other words, while the monumental task of compiling a meta-analysis of all these quantitative studies is highly interesting as an academic task, the application of the data should be taken with caution as the means by which the data has been included and standardised is at best unclear, making it somewhat questionable. Terhart is not the only person to question the statistical process in the methodology.

Arnold (2011:219) argues the 'way in which the meta-meta-analysis has been conducted..... was surprised that Hattie has chosen to summarise the effect sizes of the 800 meta-analyses using unweighted averages. Small and large meta-analyses have equal weight.....Instead I would have opted for weighted averaging by number of studies, students or effect sizes. At a minimum, it would be interesting to see whether the results are robust to the choice of averaging.' He goes on to note that he selected one study from the reference list to explore in detail (Kuncel et al, 2005) and found that the paper 'is about the validity of ex-post self-reported grades (due to imperfect storage and retrieval from memory or intentional deception), *not* about students' expectations or their predictive power of their own study performance, as Hattie claims. The paper thus should not have been included in the analysis.' While this was a single sample of 1, it does raise questions as to the rest of the studies and the lack of weighting for effect sizes.

This made me delve into the appendices of the book to review the meta-analysis themselves. In Inquiry based teaching there were 4 studies contributing, three of which applied to science, and one relating to critical thinking skills specifically. Moreover, on closer look, one paper focussed on teacher education (Sweitzer & Anderson, 1983), one focussed on the impact of assessment methods on results (Shymansky, Hedges & Woodworth, 1990), one was an unpublished thesis (Smith, 1996), and one was a paper exploring the 'newness of the method' in teaching critical thinking (Bangert Drowns, 1992). None specifically looked at the impact on inquiry based teaching in comparison to other teaching methods – the focus was on how factors impacting teaching impacted on inquiry based teaching.

This was a below average impact overall. In problem-based learning there were 8 studies contributing which are equally peculiar in the insight they can offer school teachers. Three were in the domain of medicine which is not a school based subject, but a university course, so the context of learning is completely different (Albanese & Mitchell, 1993; Smith, 2003; and Newman, 2004). One was an exploratory study rather than a meta-analysis (Newman, 2004), and another was a literature review (Albanese & Mitchell, 1993). Only one compared the teaching method in a school and that had the largest positive effect size of 37% (Haas, 2005), while two looked at the influence of assessment on problem based learning itself (Vernon & Blake, 1993; and Gijbels, Dochy, Van den Bossche & Segers, 2005).

While I do not want a methodological argument to be the focus of this paper, I think it is important to note that, as with all research, there is no perfect research question, and no perfect solution or methodology. The issues that need to be factored into the interpretation and application of Hattie's work are the lack of weighting of work, and the appropriateness of some of the studies in their inclusion. This is something I will reflect on further in the context and currency sections.

Context

Australia is relatively unique in terms of its demographics for Western countries. While it does have major cities, it also has a highly dispersed population living in remote and rural areas.

Schooling is provided, to some, through remote learning such as the School of the Air (SOTA), and boarding schools exist for rural kids to attend for their high school years simply so they get access to a broader education.

Many of these schools do not have access to specialist teaching in some of the subject areas that they would like, including maths and sciences, as teachers often choose not to locate themselves physically in these areas. Add to this the Union requirement for online teachers to be physically located in the school for which they provide online tuition, and the opportunity for subject specialist teaching diminishes further.

Anyone who has taught in a rural or remote school, or taught students through a SOTA experience will appreciate how different the experience is. Amongst other things, the life experience of these kids has been different to urban kids, so the frame of reference and application of learning is different.

The majority of papers considered in the analysis are from America, with a number being unpublished doctoral theses from US universities. Some are Canadian studies, and some it is difficult to tell merely from the publication details. The American schooling system, curriculum and provisions vary greatly from those in Australia, so it would be a reasonable question to ask how much the context of the meta-analysis considered would impact on the significance of the results.

In addition, as outlined in the methodology section, some of the studies were carried out in tertiary and higher education, so again the context in which the study was carried out would vary considerably to that of a school classroom for which the conclusions were drawn.

This does not mean that the analysis is of no value. It simply means that the impact of context needs to be considered. The problem-based learning studies in medicine demonstrated that a more narrow band of knowledge and skill were covered, but also showed the depth of learning was greater. If you are looking for a pedagogic approach to improve depth of learning, PBL may be appropriate as it has been proven to work for that purpose, but in a different context. If you were looking for broad surface learning, it would not appear appropriate. A simple effect size rating and a listing in a general ranking does not allow for context, and listings were not provided to account for different contexts or learning goals.

Assumptions

Drawing on the arguments from methodology and context above, a number of assumptions are made by people who simply take the results presented in the book at face value and quote elements of the book without consideration of the details.

Firstly they are assuming the methodology is appropriate and the methodology of every study is appropriate. Given they are all peer-reviewed, this is a relatively sound assumption to make. However, it demonstrates a lack of understanding of meta-analysis as a methodology. Meta-analysis combine other studies, and as no two studies are likely to have been carried out utilising exactly the same methods and methodology, there will be what are perhaps best described as the equivalent of 'rounding errors' in the combining of studies. If you then have these 'rounding errors' on the studies with 'rounding errors' the accuracy of the results flex somewhat.

Meta-analysis are great for establishing patterns and relative relationships between factors by summarising a range of studies in a field, but that is all they do. They do not establish causation, but many people assume they do.

Another assumption is that context doesn't matter. Most educators, in a normal conversation, would argue otherwise and argue that context does matter – and I would agree with them. However, when it comes to Hattie's work, educators seem happy to accept the results without questioning the context in which they were generated. We've already noted above that this is largely an American context.

A third assumption is that the role of the teacher and other factors not directly related to the pedagogy do not matter, despite the fact that Hattie clearly states that the teacher is the most important factor overall. By presenting a ranking of impact factors there is no cross-factor analysis. So, for example, does the role of the teacher impact more on some pedagogical approaches than others?

Fourth is the assumption that each of these learning approaches or factors occur in isolation. The reality of a classroom is they don't. Instruction, feedback, team working, peer learning, different pedagogic approaches all occur within the same lesson, sometimes simultaneously, so to look at an effect size in isolation is misguided.

When I am approaching a new area of research, I am always delighted if I can find a meta-analysis or two in the field. They give me an overview map of what research has taken place, and how the land lies.

That's it. They do not give me detailed analysis of a phenomenon on which I can make assumptions to base the rest of my research. To do that, I need to go back into looking at the individual studies that formed part of the meta-analysis – but as an overall guide to the field of study, they're great.

To review Hattie's findings on inquiry led and practice based learning, for example, and conclude from the study that, in given contexts, they lead to deeper learning, but the assessment needs to be carefully written as the impact of assessment on the process is great and would be a sound interpretation of the findings. To conclude that 'Hattie says inquiry led and PBL are not as effective as instructor led learning in schools' is more difficult. That is not what the results show, but it is what the results have been interpreted as, because the assumptions around methodology, context and role of the teacher are misplaced.

Unfortunately when you publish a study, you have no idea how people will quote, interpret and utilise your data. I have seen many reviews and references to Hattie's work that note his highlighting the importance of the teacher. This is done to boost the relevance and status of teachers, and to pat them on the back for being so important. It is generally presented as a positive outcome. What they don't do, however, is note on p250 that he laments on the poor quality of teachers overall – not quite the positive stance projected.

Teachers and Leadership

Leadership in education needs to occur at all levels. From the Government (as exemplified with New Zealand engaging Hattie to further the results of his work in their context); to the individual Schools (as exemplified by outstanding Principals who bring about wholesale change in their individual institutions); to the individual teacher in the classroom. A failure of leadership leads to a failure to thrive. Where the Principal leads, so do the teachers develop as leaders, and hence the whole school thrives. Without a Principal with a vision, some teachers may lead, but many will drift. And without strong Government direction, a whole nation's education will drift, as is evidenced by Australia today.

Perhaps the most pertinent paragraph from Hattie's (2009) book is found on p250, where he reflects overall on the experience of teaching that most children will experience:

'It is sobering to realize that we have a teaching cohort that is average, at best, in the eyes of most students. It is sobering to realize that each child will meet only a few teachers who they will consider to have a lasting and positive effect on them. It is sobering to realize that these teachers will be remembered not because they taught social studies or mathematics but because they cared about teaching the students their passion for their subject, gave students confidence in themselves as learners and as people, treated the student as a person, and instilled a love of learning of their subject(s).'

If you put his comments on teaching PBL and inquiry-based learning methods into this context, then of course, the notion of such methods being used as hands-off, unfacilitated and unstructured are expected. Good PBL teaching requires teachers to be on the ball to stay up with where their students are exploring, developing their own subject knowledge as they go, and facilitating the fun and excitement that students experience as they discover something new. We found this early in our testing of the Inventorium¹. When we train and manage the teachers, we have 100% retention rate with students. When we left it to teachers in classrooms, we pretty much had 100% non-engagement! The teacher is the most important person in learning, regardless of the methodology, and while teachers in Australia are currently taught to teach from the front of the class to a single specialist subject area, we will never succeed in shifting to the PBL, inquiry-led approach that addresses so many of the needs of the educational agenda currently.

Currency

The final point I want to make with regards to the interpretation of Hattie's work is the inevitability of time delays that occur in meta-analysis. This is partly due to the requirement for enough studies to be published that can be meta-

¹Inventorium is an online alternative high school provision that meets the Australian Curriculum Standards, but does so through an interdisciplinary, inquiry led and problem based learning approach (see <https://inventorium.com.au>).

analysed to start off with. Add to this the time taken to meta-analyse these meta-analyses. Hattie's work was published in 2009. The majority of studies included in his analysis were published in the 1980's and 1990's, and even those published in the early 2000's would have used data collected in the 1990's.

To put this in perspective, the modern usage of the internet kicked off in the 1990s. Mobile phones became widely available in the early 2000's. We're now living in a world where libraries do not rely on micro-fiche to catalogue documents, and information is instantly searchable and available everywhere.

Kids bring mobile phones into the classroom, have a laptop or tablet, and we are all contactable and traceable 24/7 if we want to be. The virtual world has made physical presence at learning events an option rather than it previously being the only option.

None of these features are reflected in the studies that Hattie's work looked at. The studies that Hattie summarises were conducted in classrooms with kids sitting in rows, learning from a teacher who used chalk, or at best a pen on a whiteboard. Indeed television is noted as having a negative effect size in the listings, while most young people don't watch television – they stream video or use social media – neither of which existed at the times of the study. If you wanted to find information, you had to go to the library. I know, because I grew up then – and the role of the teacher then was very different to the role of the teacher now.

If none of the arguments in this paper undermine your confidence in the conclusions that have been drawn from Hattie's work and how they currently apply to Australia, this one is irrefutable. Would you want a surgeon looking at books from 1990 to learn about the use of robotics in surgery? No! Would you want an oncologist to look up the latest cancer treatments from the 1980's? No! So why are you allowing teaching outcomes from this time to influence what you are doing today in the classroom? You're relying on studies from a time when the teacher was the font of all knowledge in the classroom to inform practice when the kids can check what you're saying or find more updated evidence on their phone! Being the subject matter expert is no longer the role of the teacher; teaching passion for a subject and learning is.

Hattie is also not stuck in his interpretations to a 20th century view. In his later work, Hattie and Donoghue (2016) set out to complete another ranking meta-analysis, but abandoned the ranking as they found too much variation within strategies. Instead they developed a learning model that has three inputs to learning that students bring to the process: skill or prior learning, will or disposition to learn, and thrill or motivation to learn; and three identified phases of learning: surface, deep and transfer. Amongst their conclusions is the recognition that the optimal strategies for learning depend on where in the cycle the student is, recognising that the criticisms made of PBL in the Visible Learning book could be accounted for by the fact that they were meta-analysis of first year medical students who were still at the surface learning stage, and hence PBL as not the right strategy for them at that time as students do not have sufficient knowledge to engage in forward reasoning (see also Albenese & Mitchell, 1993).

Evaluating the Contribution of Academic Discourse to Policy Development

While it is great to see policy makers and Governments adopting an evidence-based approach to policy development and implementation, it is important that they evaluate the academic offering in terms of the merits it offers the policy agenda. Simply adopting an academic argument because it is appealing and at face value seems applicable can be dangerous. We have also seen this with a rush to approve drugs for treatment of COVID19 quickly being redacted as further trials suggest that the drug does not offer the promise initially evidenced when applied on a wider scale, or indeed has side-effects which are worse than the benefits.

This paper has outlined 5 reasons why Hattie's work has not been appropriately used in Australian policy development. Had these reasons been considered at the time that policy was developed, the work would still have had value and been useful, but it would have been interpreted and used differently to the way in which it was.

These elements would be relevant in any policy context and can be presented as a conceptual framework for evaluating academic studies with regards to their impact in policy development.



Figure 1: Framework for evaluating the contribution of academic studies to policy development.

Closing Arguments

In his talk in the Lessons from Lockdown Conference (2020), Hattie talks about Covid having given us the largest educational experiment that would never have got ethics approval if proposed as a research project. He talks about curriculum not making much difference to the class as 70% of questions asked by teachers are about facts requiring 3 word answers or less, and that this is the ‘old grammar of schooling.’ He asks ‘do we really need to have all the students sitting in front of a teacher every day?’ and sees the teacher’s role shifting from involving less talking and more dialogue about the process of learning. He also notes that failure is a learner’s best friend, but this is only true if failing is low-stakes rather than the high-stakes testing regimes of NAPLAN etc. Learning through failure is part of PBL, and recovered through inquiry led learning. The learning model Hattie advocates for now is well met by PBL and inquiry led pedagogies, but it does still require teachers who are skilled at implementing the pedagogic approach.

Visible Learning is a fantastic study; it is a remarkable feat of statistical analysis; and it is a good map of how teaching practices impacted on learning outcomes around the 1980-1990s, before the use and possibility of technology in the classroom became all but ubiquitous². It is a book that looks at Western education practice and notes the importance of the teacher, reflecting also on the general poor quality of teaching that was occurring at that time.

These issues are systemic and need leadership in Government in order for them to be addressed at a systemic level, else the bureaucratic processes that hold them in place will prevent them from changing. Teacher training needs to be rethought to support the development of teachers for the current century’s learning model, and professional standards need to be rethought to encourage innovative practices and a true love of teaching and learning, rather than adherence to process.

Australia has lost its way in education and Hattie is the fallback that is used to justify this position. But I doubt this was Hattie’s intent. Indeed, he went on to lead great work in New Zealand’s education system, applying the lessons he learnt from his work in their specific context, and has since moved from ranking learning strategies to best fitting them within a learning model (Hattie and Donoghue, 2016).

² We do note that the internet is not fully accessible easily everywhere, but this is viewed as an infrastructure issue for fixing rather than an educational issue. Until it is resolved, downloads etc need to be utilised, or alternates produced to tech use in learning, but this will be a short-term requirement only.

Australia, on the other hand, read it as a good excuse not to change and so here we are, over a decade on, still citing evidence drawn from the last century to drive our behaviour in schools. Is that what Hattie intended? Not if you listen to what he's saying today, or actually read the book fully.

References

- Albanese, M. A., & Mitchell, S. (1993). Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68(1), 52–81.
- Arnold, I. & Hattie, J. (2011) Visible learning: A synthesis of over 800 meta-analyses relating to achievement. *Int Rev Educ* 57, 219 <https://doi.org/10.1007/s11159-011-9198-8>
- Bangert-Drowns, R. L. (1992). Meta-analysis of the effects of inquiry-based instruction on critical thinking. *Paper presented at the Annual meeting of the American Educational Research Association*, San Francisco, CA.
- Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. (2003). Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, 13(5), 533–568.
- 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(1), 27–61.
- Haas, M. (2005). Teaching methods for secondary algebra: A meta-analysis of findings. *NASSP Bulletin*, 89(642), 24–46.
- Hattie, J. A. C. (2009) *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge, Oxfordshire.
- Hattie, J. A. C. & Donoghue, G. M. (2016) Learning Strategies: A synthesis and conceptual model. *Science of Learning*. 1: 16013
- Hattie, J. (2020) Lessons from the Lockdown Conference. 1 Sept. @ <https://lessonsfromthelockdown.com>
- Kuncel, N. R., Crede, M., & Thomas, L. L. (2005) The validity of self-reported grade point averages, class ranks and test scores: A meta-analysis and review of the literature. *Review of Educational Research*, 75(1), 63–82.
- Newman, M. (2004). *Problem-based learning: An exploration of the method and evaluation of its effectiveness in a continuing nursing education programme*. London: Middlesex University.
- Shymansky, J. A., Hedges, L. V., & Woodworth, G. (1990). A reassessment of the effects of inquirybased science curricula of the 60's on student performance. *Journal of Research in Science Teaching*, 27(2), 127–144.
- Smith, D. A. (1996). A meta-analysis of student outcomes attributable to the teaching of science as inquiry as compared to traditional methodology. *Unpublished Ed.D.*, Temple University, Pennsylvania, United States.
- Smith, R. A. (2003). Problem-based versus lecture-based medical teaching and learning: A meta-analysis of cognitive and noncognitive outcomes. *Unpublished Ph.D.*, University of Florida, FL.
- Sweitzer, G. L., & Anderson, R. D. (1983). A meta-analysis of research on science teacher education practices associated with inquiry strategy. *Journal of Research in Science Teaching*, 20(5), 453–466.
- Terhart, E. (2011) Has John Hattie really found the holy grail of research on teaching? An extended review of Visible Learning. *J. Curriculum Studies*, 43(3), 425–438.
- Vernon, D. T., & Blake, R. L. (1993). Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine*, 68(7), 550–563.
- Walker, A. E., & Leary, H. (2008). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem Based Learning*. 3(1).