

ASSESSING CLAIMS IN EDUCATION

CLAIMS

Statements which present themselves as being true, but the truth of which we cannot be certain, are called claims. It is important to consider how to treat such claims and how far they can be trusted. In order to do so, you need to analyse the thinking behind such claims.

Beware of claims that are “too good to be true”

It is more compelling to hear of an educational intervention that will increase outcomes by 80% rather than 5%, or that is “guaranteed to work” – but often claims which make bold promises are precisely the sorts of claims to be sceptical of. Whenever you are presented with a claim that promises ‘certain’ results, or huge improvements, remember that large, dramatic effects of interventions are rare, and interventions can cause harms as well as benefits.

Amelia is reading about ‘Neuro-Move’, an intervention which teaches students a number of simple physical activities to be performed during lessons. The intervention claims these exercises ‘integrate body and mind’ and bring about ‘rapid and dramatic improvements’ across a range of outcomes, including concentration, memory, reading, writing, organising, listening, physical coordination, and more. On the basis that it ‘sounds too good to be true’ she decides to look more closely into the evidence before considering the intervention for her school.

Check that claims are based on sound logic – watch out for faulty assumptions

Although professional judgement and beliefs are an important and inevitable factor when making decisions, they are not, by themselves, a reliable predictor of impact. Whilst there may be some logic to support a claim, remember that some ways of thinking – while psychologically persuasive – are nevertheless flawed. There are lots of ways in which we are prone to making biased judgments. Watch out for faulty assumptions such as: assuming that just because something is new it is bound to be good; or conversely, that because an approach or product is widely used or well-established it must be beneficial. Likewise, don’t assume that increasing the amount of an intervention or strategy will necessarily increase the impact – there may be lost opportunities in overlooking alternative, potentially more beneficial, approaches.

Yusuf’s school encourages teachers to match lesson activities to a pupil’s preferred ‘learning style’. When he asks about it, he is told that lots of teachers believe it to be effective and the school has planned lessons in this way for many years. On the basis that ‘we’ve always done it this way’ isn’t a reliable basis for judging effectiveness of an intervention, Yusuf decides to investigate the evidence.

Another common error is to assume that just because something is associated with a particular outcome it caused it to occur. It is entirely possible that the relationship is coincidental. Identifying the causal impact of an intervention depends on making fair comparisons with alternative approaches, including not doing the intervention (see Comparisons below).

Zainab and Owen are colleagues within the same school. The school has recently had an 'aspirations week' focused on extra-curricular activities intended to help pupils develop the attitudes and behaviours to support them achieving well in school. Owen believes the intervention to be a success, as following the intervention there was an increase in school attendance and a reduction in lateness. Zainab is less certain, arguing that these changes might be due to other factors, including recent parents' evenings and the start of revision sessions for year groups doing exams. She argues that a fairer comparison would need to be made to judge whether 'aspirations week' caused these changes.

Check that claims are based on more than personal experience

The fact that someone else believes in an intervention, or reports to have had a good experience with it, does not necessarily mean it is better than an alternative. We have a natural tendency to overestimate the effects of the interventions that we invest time, effort and money in, meaning personal anecdotes and experiences are not necessarily a reliable basis for judging their impact.

Watch out for claims where the person stating the claim may have a vested interest in you believing it, such as a financial interest in an education programme. Even if an intervention is endorsed by an expert or authority figure, don't rely solely on their opinion – look for evidence from fair comparisons to support that claim (see Comparisons below).

Naimh is a head teacher for a medium sized primary school. Over the holidays she read a book by a well-known professor of education which argued that schools are failing to keep up with the changing demands of the 21st century and that providing a laptop to every pupil would help boost pupil attainment. Naimh notes that the book has been endorsed a large company which makes and sells laptops to schools. She decides to read more independent evidence on the costs and benefits of using digital technology before recommending the approach to her governing body.

COMPARISONS

Different types of research are useful for different purposes. For example, a small-scale qualitative study, which uses surveys and classroom observations, can be helpful in understanding how or why a particular approach works. Conversely, the most reliable way of establishing the impact of an intervention, or validity of a claim, is to compare it to an alternative intervention or not intervening at all, using an experimental trial. Nevertheless,

just because some form of comparison has taken place does not necessarily mean it is a good one. Some 'evidence', just like some ways of thinking, can be faulty and unreliable.

Don't be misled by unfair comparisons

As discussed, not all comparisons are fair comparisons. Typically, a fair comparison involves keeping as many things as consistent as possible – the conditions, the type of participants, the measures etc – apart from the interventions being compared. Beware of comparisons where the groups being compared are treated or assessed differently (aside from the intervention), or where outcomes are collected from only a proportion of participants in a study.

Elias is at a TeachMeet where colleagues are presenting findings of some action research conducted in their schools. In one presentation, the teacher explains a mentoring intervention they trialled with their classes and presents results which appear to show significant gains in comparison to another class. Elias wonders how fair this comparison was and asks the teacher after the presentation to explain more. It becomes apparent that the group receiving the intervention was a 'top set' within the school and mainly girls, and that the comparison group had lower prior attainment, and were predominantly boys. As a consequence, Elias remains sceptical as to whether the intervention genuinely had the impact the teacher suggested.

Don't be misled by unreliable summaries of studies

Where possible, evidence that is aggregated from more than one evaluation, in a review, will be more reliable than evidence that is drawn from a single study. Reviews that use systematic methods to reduce bias, and increase transparency on how data is analysed and interpreted, are more likely to be reliable than reviews that don't.

Ester is reading a blog post which claims to summarise the evidence on the best ways to teach reading. She notices, however, that the author only appears to provide descriptions of a small number of studies which support their favoured approach, and neglects other studies which present contrasting evidence. She decides to search for a more systematic summary of the evidence before forming an opinion.

Don't be misled by how effects are described

Wherever possible, look for data to support a claim for an intervention, rather than just a verbal description of the effects that were observed. Words like 'dramatic' or 'sizeable' have no precise meaning and are open to interpretation (who determines what actually constitutes a 'dramatic' improvement?).

At the same time, the way in which data is collected, analysed and reported can also be misleading. Beware of studies with small sample sizes, which will typically have less reliable findings. Likewise, the idea of a % improvement can be very misleading. "100% increase" could mean an increase from 1 – 2%! Presenting confidence intervals for a study/comparison provides a useful way of determining the likelihood that the result

could have occurred by chance, although beware of confusing ‘statistical significance’ with ‘importance’. Statistically non-significant results can also provide useful information i.e. showing that it is unlikely that alternative interventions differed in their effects.

Bodhi is head of assessment in his school. At the beginning of the school year he receives data comparing school results to national averages. At first he is excited to see that maths attainment appears to be above average this year – especially as they had introduced additional homework clubs over the year. However, on closer inspection he can see that the confidence interval for the school result crosses the line for national average, and that the analysis suggests that the school’s result is not significantly different from national average statistically. However, whilst not statistically significant, the result still represents interesting data – as it’s possible the investment homework clubs didn’t produce the impact they expected.

Finally, recognise that lack of evidence is different to evidence of ‘no difference’. There are lots of many interventions and strategies in education that haven’t been rigorously evaluated using fair comparisons, meaning we know little about their effects.

CHOICES

Even the most reliable evidence for an intervention or claim might not be directly applicable to your context. You may be facing a different problem, or may want to achieve another goal. Your students may be different from the participants in a study in terms of their age or socio-economic background, or the intervention might simply be too expensive to be rolled out on a larger scale. It is important to consider a range of factors and questions before deciding whether or not a given intervention is suitable for your context.

What is the problem (or what are the goals) and what are the options?

Before embarking on any course of action check that the problem has been diagnosed correctly and the goals are appropriate. Is the issue that you are trying to improve clear? Do you know exactly what it is that you are hoping to achieve? Is the intervention well-placed to achieve those objectives? It is important that the options being considered include all of the relevant ones available to you, not necessarily just the ones included in a given study.

Nansi is leading a discussion with her senior leadership team regarding ‘wellbeing’ within the school. One of the points emerging quickly in the meeting is that members of the team have a different understanding of the term – some relating it to emotional states like ‘happiness’ or ‘anxiety’, others relating it to physical health like ‘sleep’ and ‘physical activity’, others consider it a term related to mental attributions like ‘grit’ or ‘positivity’. She decides that before considering possible interventions the team need to have a clear, shared understanding of what the problem is and what they want to achieve.

Is the available evidence of the effects relevant?

Although the evidence for a claim or intervention might appear to be solid or impressive, the approach might not necessarily transfer well to your context. It may be, for example, that the people in the studies are very different to those of interest to you, or that the circumstances in which the interventions were compared are very different from your own. If this the case, the described effects may not be applicable or transferable to your context.

Joel is a secondary head teacher and has been reading research suggesting potential benefits for disadvantaged pupils when schools started and finished an hour later in the day. He notes, however, that the studies are almost exclusively based in the US rather than the UK, where schools typically start earlier. Joel questions whether the apparently positive results would apply given the contextual and cultural differences between these schools and his own, and decides to look for any studies based in the UK before proposing any changes to the timetable.

Do the expected benefits and savings outweigh the expected harms and costs?

Perhaps the most important question to ask! Very few interventions have purely positive outcomes: often there is a cost on some level, even if it's just time that could have been spent on alternative courses of action. As well as considering the possible benefits, also consider: the likelihood of the benefits happening; any potential harms; how important the benefits are to you (and how undesirable the potential harms may be); and what the costs are in terms of time, effort and money (including things you will not be able to do in its place).

Tilly is head of English in a secondary school. Over the year, the department has been trialling a new marking strategy in the hope that the improved feedback given to pupils would help raise attainment. However, she has noted some things which suggest the policy may be having some negative impact – including members of staff staying late to complete marking and the cancellation of an after-school creative writing club because the staff member no longer had time to run it. She decides it would be worth reviewing the marking policy of the department to bear into consideration these costs as well as the potential benefits.

