
**While Heisenberg is not looking: the strength of ‘weak measurements’ in educational research**

David R. Geelan  
Griffith University  
d.geelan@griffith.edu.au

**Abstract**

The concept of ‘weak measurements’ in quantum physics is a way of ‘cheating’ the Uncertainty Principle. Heisenberg stated (and 85 years of experiments have demonstrated) that it is impossible to know both the position and momentum of a particle with arbitrary precision. More precise measurements of one decrease the precision with which the other can be measured. By ‘sneaking a peak’ at one variable however – conducting a ‘weak measurement’ that does not fully collapse the quantum wavefunction – and combining this with a subsequent strong measurement of the other variable, increased information of better quality can be gathered. By analogy, theories and methodological approaches could conceivably be combined in research in such a way that one does not ‘force [the other] into blindness’ (Bauersfeld 1988). This paper offers a theoretical and metaphorical introduction to an approach. The criteria for judging its value are empirical: does it work? Does it allow research work to be done and conceptualised in ways that are valuable? That evidence will be forthcoming as this approach is adopted and tested.

**Keywords:** Theory, Methodology, Eclecticism, Uncertainty

**Introduction**

In a number of previous discussions of the ways in which methods and theoretical perspectives are combined in educational research (Geelan 1997, 2001, 2004) I have mentioned, but slid away from, the issue of ‘incommensurability’. The word means, literally, ‘not measurable on the same scale’ – like love and gold, for example.

Theoretical perspectives in positivist and post-positivist traditions, similarly, may be founded in quite different fundamental epistemologies, and there are questions about the ways in which it might be valid (or otherwise) to combine them in seeking to understand educational questions. Similarly, research methodologies from the qualitative and quantitative traditions, broadly described, or from much closer cousins such as phenomenology (van Manen 1990) and phenomenography (Marton, 1988), may focus attention on different aspects of educational situations in ways that raise questions about the wisdom and effectiveness – or indeed the legitimacy – of combining them.

In the past (Geelan, 2001) I have used perspectives from the philosophy of science, including those of Kuhn (1970), Popper (1965, 1968), Lakatos (1970) and Feyerabend (1974, 1975), as a way of exploring some of these issues in relation to combining theories. This is not a matter of a direct application of the philosophy of science to questions in educational research, but the use of such
perspectives as a set of valuable metaphors to be used as tools for thinking through the problems that perplex us. The physical sciences and the associated philosophy are taken, not as ‘higher’ sources of truth or validation, but simply as sources of interesting and potentially valuable metaphors.

This paper takes some recent work in the field of quantum physics (Dressel et al., 2013; Kocsis et al., 2011; Lundeen et al., 2011) and uses it metaphorically to aid readers (and myself) in thinking through some issues in relation to how theories and methodologies can be productively combined in educational research.

This discussion is, of course, part of a very long conversation in educational research about these issues. As early as 1989, Nathaniel Gage (1989) published a (n imagined) ‘historical’ overview of the ‘paradigm wars’ in education and their aftermath, suggesting, if not peace, at least an armed detente between the qualitative and quantitative camps. Tashakkori and Teddlie (2003) have been working thoughtfully in the fields of ‘mixed methods’ research, and a recent paper by Hall (2012), drawing on a term from Kuhn (1970), addresses the perception that the field is still ‘in search of a paradigm’.

Heisenberg’s Uncertainty Principle, while it has been used (and abused) as a metaphor for many things, has been used only fleetingly (Bauersfeld, 1988; Geelan, 1997) to think about combining theories and methodologies in educational research. This paper takes that metaphor, along with more recent work on the concept of ‘weak measurements’ (Dressel et al. 2013; Kocsis et al. 2011; Lundeen et al. 2011), to explore some novel approaches to addressing these issues.

The Uncertainty Principle

In 1927, in a paper entitled ‘Ueber den anschaulichen Inhalt der quantentheoretischen Kinematik and Mechanik’, Werner Heisenberg stated what he called the ‘indeterminacy relation’, but has come to be referred to as the ‘uncertainty principle’. It places a boundary on the precision to which the position and momentum of a particle can be simultaneously measured, and seems (based on all the empirical evidence available to date) to be a fundamental feature of the physical laws of the universe rather than an artefact of flaws in our own instruments and abilities to make measurements.

In practice, since the two uncertainties are multiplied and there is a limit placed on their product, decreasing the uncertainty in one variable increases the uncertainty in the other, and vice versa. We can know either the position of the particle or its momentum (which is proportional to its speed) as precisely as we like, but the more precisely we know one, the less we know about the other. If we can completely specify the position, we can know nothing about the momentum, and vice versa.

Over a number of years I have been considering the ways in which theories and methodologies that have different underlying assumptions, and may be incompatible or even incommensurable, can be combined in educational research (Geelan 1997, 2001, 2003, 2004). With Bauersfeld (1988) I have noted that ‘Concentrated and systematic pursuit of one perspective forces other perspectives into blindness’ (p. 41). This paper explores the use of ‘weak measurements’ in physics as a metaphor for thinking more deeply about these issues of theoretical and methodological fluency and the ways in which eclectic multimethod research is conducted.

Weak measurement in physics
The question of ‘measurement’ is a vexed one in quantum mechanics – what constitutes a ‘measurement’? Plenty has been written on the topic (Busch et al., 1996, provide a good overview) and in general it reduces to the notion that a ‘measurement’ moves some ‘pointer’. A ‘weak measurement’ occurs when the strength of the coupling between the physical system being measured and the pointer is reduced. The weaker coupling means that the measurement makes a smaller impact on the quantity being measured. It also reduces the precision of the measurement, but averaging many measurements can compensate for this reduced precision. As the strength of the coupling approaches zero, the perturbation also approaches zero... but the precision of the measurement decreases.

Two recent papers in physics (Kocsis et al., 2011; Lundeen et al., 2011) have demonstrated the use of ‘weak measurements’. While the details of the two studies vary, the key features are the same: the use of ‘pointers’ that are only weakly coupled to the first variable to be measured, followed by ‘post-selection’ on the other variable. The Lundeen paper describes an experiment in which position was measured weakly followed by a strong measurement of momentum, while the Kocsis paper describes weak momentum measurements followed by strong position measurements, so together they offer a nice complementary (in a non-quantum sense) set of demonstrations of the possibilities.

Heisenberg’s principle states that for certain variables – position and momentum represent one such set, but there are others – measurements on a particular particle are linked. The key to ‘weak measurements’ is to, if you will, not measure the position of the particle but... sneak a peek at it out of the corner of our eye between our slitted fingers, so that we have an approximate idea of its position. Since we do not know precisely where it is, but only a range of possible values, the quantum wavefunction is not fully collapsed, only partly.

In the actual research studies (Kocsis et al., 2011; Lundeen et al., 2011) the second variable of interest is then ‘post-selected’: if position was measured weakly, all the particles with a particular momentum are selected at the end of the experiment, and only the weak measurements from those particles are used. It is the combination of weak measurement and post-selection that allows the Uncertainty Principle to be, not circumvented, but mitigated.

The notion that two (or more, but for the purposes of this discussion I’ll keep it at two) theoretical perspectives or methodological approaches can be fruitfully combined is something I have written about explicitly in the past (Geelan 1997, 2001). The remainder of this paper applies the metaphor of ‘weak measurements’ in quantum mechanics to combining theories and methodologies in education. It is worth noting that ‘weak’ in this context does not mean ‘bad’ or ‘of poor quality’, and is not any kind of negative judgement. On the contrary, weak measurement is a tool in experimental quantum mechanics that allows science to explore areas that are impossible to access with traditional approaches (Dressel et al., 2013).

Given some of the common metaphors in the field of educational research, it may be tempting to assume that ‘weak’ measurements correspond to qualitative, postpositivist or naturalistic methods and ‘strong’ measurements to quantitative, positivist or experimental methods, but this would be to misunderstand. It would be possible to measure the quantitative features of an educational context of interest ‘weakly’ and to combine that with a ‘strong’ qualitative measurement of the same context, for example.

It is hoped that the notion of ‘weak measurements’ in educational research will contribute to addressing issues of educational interest in genuinely novel ways.
Combining theories in educational research

More than 17 years ago I wrote:

By analogy with Heisenberg’s Principle “one cannot elaborate two incompatible perspectives with arbitrary precision. Concentrated and systematic pursuit of one perspective forces other perspectives into blindness” (Bauersfeld 1988, p. 41)..... By alternating between (possibly incompatible) perspectives, however, I would suggest that it is possible to see into some of our blind spots. (Geelan 1997, p. 24)

At that time, and subsequently (Geelan 2001), I suggested a ‘disciplined eclecticism’: the careful, thoughtful selection of two or more different theoretical perspectives used in juxtaposition to allow each to highlight features that the other ‘forced into blindness’. I elided the issue of the potential incommensurability of theoretical perspectives. I suggested that in a practical sense, in a context where judgements about the value of educational research are made using standards of verisimilitude (seeming true to those in a position to make that judgement) and utility (usefulness for the purposes of researchers and practitioners in the field, generativity for prompting new investigations), it didn’t really matter that incommensurable perspectives were being combined.

This approach was never fully intellectually satisfying to me, and later – half in the interests of provocation and half in earnest – I suggested that educational research dispense with theory entirely (Geelan 2003). Eschewing theory has been proposed by others, including Thomas (2007). I later resiled from this position, reviving the ‘undead theories’ in a collection of my papers (Geelan 2006).

Heinrich Baursfeld (1988) cautioned in relation to combining theories:

...one cannot elaborate two incompatible perspectives with arbitrary precision. Concentrated and systematic pursuit of one perspective forces other perspectives into blindness. ... The need for coherent description and explanations (or interpretations) then makes the other perspective reappear as 'noise', as 'complementary issues', as 'intervening variables' and so forth, which hamper the precision of description and weaken the predictive power of the theory. (p. 41)

I do think there is value in thinking through the implications of ‘incommensurability’: the notion that theories describe different things on different scales that cannot be unproblematically synthesised or combined. Positivist approaches have quite different assumptions about the roles and purposes of research and education, and about what ‘counts’, compared to post-positivist and postmodern ones, and there is scope for a deeper conversation about what we are doing when we combine these theoretical perspectives in an eclectic way.

I described a dialectical approach to the use of incommensurable theoretical perspectives:

To think dialectically in a research situation, then, is to metaphorically put on the ‘spectacles’ provided by one theoretical perspective, learn what can be learned of the situations, contexts and events which are of interest, then remove those spectacles and replace them with a different pair (in some ways, the more different the better), and see how the view changes.
What is visible now that was hidden before, and vice versa? Since all observations are of necessity theory laden, the most powerful approach available for increasing the usefulness of observations is to change the theories with which they are laden: there is no God’s eye, unbespectacled view available. ... This image is, of course, over-simple: it assumes that the ‘object’ of study will sit still under different theoretical perspectives. But the object itself changes with the theoretical perspective, as different facets of school life are selected and neglected. (Geelan, 2001, p. 141)

There are other, perhaps complementary, approaches to issues of theory, eclecticism and research.

To return to the metaphor of ‘weak measurement’, the notion of conducting and averaging a range of ‘weak measurements’ of a research situation of interest using one theoretical framework, then looking again at the same ‘object’ [though, as Bauersfeld (1988) has noted, it can never be exactly the same object, since observation in some senses ‘constructs’ the object of research scrutiny] using a different perspective remains compelling. The goal is not greater precision: by analogy with the Uncertainty Principle, greater precision (smaller uncertainty) for one perspective will yield decreased precision (larger uncertainty) for the other. Rather, a richer, more well-rounded, more sophisticated representation of the phenomena and situations of interest is the goal.

In quantum mechanics, variables do tend to come in neatly opposed dichotomous sets – such as position and momentum – which has meant that much of the discussion thus far has been about combining two theoretical perspectives. Educational situations, however, are more complex, and it is possible that a third perspective might yield, again, not greater precision but greater richness in exploring and understanding an educational situation. This richness is likely to enhance both the utility and the verisimilitude (Geelan 2004) of the research for the practitioners and others whom it is intended to serve. Decisions about adding fourth or subsequent perspectives are matters more of professional judgement than of prescription: what is gained, what is lost, what are the costs and benefits?

The issues with combining incompatible or incommensurable theories are to some extent mirrored – with some differences – in discussions around combining methodological approaches.

**Combining methodologies in educational research**

Jenny Symonds and Stephen Gorard, in their paper ‘The Death of Mixed Methods: Research Labels and Their Casualties’ (2008), suggest that the creation of ‘mixed methods’ is essentially built on stereotypical labels. ‘Quantitative’ and ‘qualitative’ are labels that combine huge ranges of quite disparate things, and Symonds and Gorard’s thesis is that these labels are unhelpful, and that their use has led to assumptions about the nature of research in education that have not well served practice in education. In particular, they argue that, precisely because the incommensurability of qualitative and quantitative methodologies has been overstated, the foundations of ‘mixed methods’ as an approach are flawed.

Citing the value of ‘triangulation’ and the complementary strengths of different methodologies, the proponents of ‘mixed methods’ ground their claims in the value of different approaches for serving the profession, but do not build a strong philosophical case. Symonds and Gorard (2008) suggest that Johnson and Onwuegbutie (2004) and Tashakkori and Teddlie (2003) have not solved the
‘incommensurability’ (or ‘incompatibility’) problem in relation to methodologies that was well known in the 1970s and 1980s, but have instead sidestepped it with an appeal to ‘pragmatism’ and ‘epistemological tolerance’. These approaches are perhaps akin to the ‘disciplined eclecticism’ I have advocated in earlier discussions of these issues (Geelan, 1997, 2001, 2004).

Symonds and Gorard (2008) claim that, since it is defined in terms of combining ‘quantitative’ and ‘qualitative’ methods, far from moving beyond such labels, it can tend to reinforce them. Their proposed solution is to move beyond labels:

A ‘radical’ suggestion is to stop trying to promote a certain type of quality through the construction of overarching categories and researcher identities, and focus rather on the quality of our actual research techniques, the resulting data and on how that data is used, no matter whether this involves one or more sets or types. (p. 17)

While I do have some sympathy for such an approach, it seems to me that, in order to be able to communicate, language and labels are essential. We may consider our labels to be useful rather than true (Polkinghorne, 1992) and to hold our methodological tools loosely, but being able to describe clearly the traditions from which we are drawing and the methods we are choosing (and our reasons for making those choices and not others) seems to me to be essential to a productive framing of the activities of research in education.

The work of those currently working in the field of ‘mixed methods’ such as Tashakkori and Teddlie (2003) and Johnson and Onwuegbuzie (2004) offers one important way to advance this project, and Symonds and Gorard’s (2008) paper offers a timely reminder to treat our boundaries as permeable rather than rigid. The notion of ‘weak measurements’ may offer a further, complementary approach.

**Weak measurement in educational research**

To reiterate, ‘weak measurements’ are not bad measurements, yielding poor quality information. Rather, they are careful, high quality measurements that perturb the phenomena being studied to a lesser degree. They are measurements that trade off precision for accuracy: precision can be reclaimed by combining multiple weak measurements.

Quantitative research in education tends to seek to either perturb the educational practices of interest not at all – in descriptive studies that seek to ‘photograph’ phenomena without the fact that the research is being conducted having any influence at all on the phenomena – or to perturb the practices completely – in experimental and quasi-experimental studies involving an intervention and a control group.

Qualitative studies, broadly defined, tend to have a more reflexive focus on the entanglements of researcher and researched, and a level of comfort – or at least resignation – toward the impossibility of conducting research in an educational context that captures the phenomena of interest as they would have been if the research had not occurred. The inability to specify in detail the extent to which and the ways in which the research influenced what was researched is accepted as a cost of doing qualitative research, a trade-off for the richer and more ‘naturalistic’ portrayals of the phenomena that such methods are claimed to make possible.
Many of the (often heated) discussions in the ‘paradigm wars’ were about a perceived lack of ‘purity’ – in the sense of not leaving the researchers’ ‘fingerprints’ on the data – on the part of some qualitative methodologies. Patti Lather (1986) embraces this, and describes the ‘catalytic validity’ of research that becomes in itself, and intentionally, an occasion for change in the very practices being studied. This approach, too, is attractive to me, but may lead to a reduction of the extent to which the research can be valuable in locations more or less unlike the research location.

It is important to reiterate that, within the context of these brief thumbnail descriptions, I am emphatically not identifying qualitative research with ‘weak’ and quantitative approaches with ‘strong’ measurements. Just as, in quantum physics, either the position or the momentum can be measured weakly and the other post-selected for strong measurement, so in education it is possible to conduct ‘weak’ quantitative measurements and ‘strong’ qualitative ones, or to have both measurements be qualitative or quantitative, one weak and one strong.

An example from my own research field – educational technologies in science education – may be helpful in outlining the promise of the ‘weak measurements’ approach. The phenomenon of interest is the use of ‘interactive simulations’ – computer-based ‘virtual laboratories’ – to support the learning of chemistry and physics concepts by senior secondary school students.

There are a number of sub phenomena, including the extent to which the interactive simulations enhance student learning, compared with other learning activities, the pedagogical decision-making and actions of the teachers, the ways in which students engage with the simulations and a range of other factors.

In this instance – the approach will vary for studies in different contexts – it is possible to capture and log data from within the interactive simulations about the ways in which students use them. The time for which students look at each page, their navigation through the site, the data they enter, the results they receive and the ‘virtual experiments’ they conduct can all be tracked. Apart from the necessity of obtaining their informed consent to participate in the research, collecting these data has no impact on the teaching and learning in these classrooms. It is a ‘weak measurement’ (that happens to be quantitative), in that it does not perturb the phenomena being studied. It is also ‘weak’ in the sense that it does not yield great precision: considerable inference is required to develop an understanding of a student’s mental journey from these kinds of quite abstracted data.

This is where the ‘post-selection’ feature of the weak measurement model becomes relevant. By conducting a post-test of students’ conceptual understanding, it is possible to determine which students have developed rich, appropriate mental models of the physics and chemistry concepts of interest. Such students can be ‘post-selected’ – as can those who developed weak or non-existent understanding – and the ‘weak’ data from the data logging can then be interpreted in the light of the information about students’ conceptual development. Other ‘strong’ (qualitative) measurements such as interviews with teachers and students, observation of lessons, or ‘think-aloud’ protocols while working with interactive simulations can also be used to ‘post-select’ students, or to complement the other available measurements.

In some ways it would be desirable to use a pre-test of conceptual understanding prior to the students’ use of the interactive simulations in order to be able to make stronger claims about their conceptual development while using the tools, but such a test would not constitute a ‘weak measurement’, since in being asked to answer certain conceptual questions the students’ conceptual...
understanding would be changed. Even these ‘objective’ quantitative data would have the researchers’ fingerprints all over them. With the data logging, however, the weak measurements, combined and post-selected, can overcome their traded-off precision to offer enhanced accuracy.

Contrary to the assumption that quantitative data are ‘strong’ and qualitative data ‘weak’, Taylor and Raykov (2014) conducted a mixed-methods study of ‘community service learning’ across several campuses of the University of Alberta in Canada. Participants in the study completed a course involving community service learning. Quantitative data in relation to participation and completion of the course by students were collected (‘weak measurements’, in the sense that they form weakly-coupled proxies for students’ actual levels of engagement with service learning) and followed up by online surveys (generating both quantitative and qualitative data) as well as interviews and focus group discussions. Since they more explicitly and directly addressed the phenomena of interest, these qualitative data would be considered the ‘strong measurements’ in this mixed-method study. Based on these ‘post-selected’ measurements it was possible for the authors to revisit the ‘weaker’ quantitative measurements conducted during the teaching of the courses.

Conclusion

The approach discussed in this paper for research in education is tentative, suggestive. It is drawn metaphorically from quantum physics but perhaps involves levels of uncertainty significantly greater than those in Heisenberg’s (1927) principle. It addresses issues of the incommensurability of different theoretical and methodological approaches combined in the same study, and offers the potential to more richly explore blended approaches to both theory and methodology.

The proof of the pudding is in the eating. In speaking of psychology, Donald Polkinghorne said “[t]he criterion for the acceptability of a knowledge claim is the fruitfulness of its implementation” (Polkinghorne, 1992, p. 162). Bauersfeld said “In the end, it is the life in mathematics classrooms which processes and provides the evaluation for theories in mathematics education” (1988, p. 42). The evidence to test a physics theorem arises in the laboratory, and the utility and value of the perspective described in this paper will only become apparent as it is adopted, used and tested in educational research.

References


