A quantum measurement paradigm for educational predicates: Implications for validity in educational measurement

Cantley, I. (2017). A quantum measurement paradigm for educational predicates: Implications for validity in educational measurement. *Educational Philosophy and Theory*, *49*(4), 405-421. https://doi.org/10.1080/00131857.2015.1048668

Published in:
Educational Philosophy and Theory

Document Version:
Peer reviewed version

A Quantum Measurement Paradigm for Educational Predicates: Implications for Validity in Educational Measurement

Ian Cantley

Queen's University Belfast

Biography

IAN CANTLEY is a lecturer in education (mathematics) at Queen's University Belfast. His research interests include the philosophy of education, the mathematical and philosophical foundations of educational measurement models and mathematics education. ADDRESS: School of Education, Queen's University Belfast, 69-71 University Street, Belfast, UK, BT7 1HL. Email: [i.cantley@qub.ac.uk]

Abstract

The outcomes of educational assessments undoubtedly have real implications for students, teachers, schools and education in the widest sense. Assessment results are, for example, used to award qualifications that determine future educational or vocational pathways of students. The results obtained by students in assessments are also used to gauge individual teacher quality, to hold schools to account for the standards achieved by their students, and to compare international education systems. Given the current high-stakes nature of educational assessment, it is imperative that the measurement practices involved have stable philosophical foundations. However, this paper casts doubt on the theoretical underpinnings of contemporary educational measurement models. Aspects of Wittgenstein's later philosophy and Bohr's philosophy of quantum theory are used to argue that a quantum theoretical rather than a Newtonian model is appropriate for educational measurement, and the associated implications for the concept of validity are elucidated. Whilst it is acknowledged that the transition to a quantum theoretical framework would not lead to the demise of educational assessment, it is argued that, where practical, current high-stakes assessments should be reformed to become as 'low-stakes' as possible. The paper also undermines some of the pro high-stakes testing rhetoric that has a tendency to afflict education.

Keywords

educational measurement, Wittgenstein, Bohr, Newtonian, quantum-theoretical, validity

Introduction

This paper challenges the theoretical foundations of contemporary educational measurement models and proposes an alternative paradigm for the measurement of educational predicates. It is argued that current educational measurement practices are based upon a Newtonian conception of measurement. The quantum pioneer Niels Bohr was adamant, however, that the properties of intentional psychological predicates (such as learning) fit the framework of quantum theory, which superseded Newtonian mechanics (that functions perfectly well for macrosystems) as the best physical model of microsystems in the early part of the twentieth century. Although Bohr referred to structural parallels between quantum theory and psychology (and, by extension, education), he never developed these claims; they are, however, investigated in the current paper. Ludwig Wittgenstein's later philosophy contains an extensive analysis of the nature of intentional predicates and this is used, together with Bohr's philosophy of quantum mechanics, to demonstrate that a quantum theoretical paradigm would, in fact, provide a more secure basis for educational measurement. Finally, the implications of a transition from a Newtonian to a quantum theoretical framework for the concept of validity in educational measurement are considered.

Psychology's "physics envy"

From the advent of psychology, some psychologists have sought to have the discipline recognised as a quantitative science and, according to Lagemann (2000), educational researchers simply followed suit. In formulating quantitative theories, psychologists re-defined measurement to suit their own purposes and attempted to measure psychological attributes, such as cognitive abilities, in the same way as the properties of macroscopic bodies are measured in Newtonian physics (Bruner, 1990). For example, item response theory is used in conjunction with structural equation modelling to estimate ability levels of individuals from their responses to test items. In this approach, it is assumed that ability levels can be abstracted away from the measuring instrument in the same way as a measurement of a dynamic attribute of a macroscopic body in Newtonian mechanics.

Michell (1997) casts considerable doubt on the legitimacy of contemporary approaches to psychological measurement. He asserts that, for measurements of psychological attributes to be legitimate, what he refers to as the "scientific task" (Michell, 1997, p. 359) of measurement must be properly investigated prior to commencing work on the "instrumental task" (Michell, 1997, p. 359). In other words, it is vital to demonstrate that psychological attributes are quantitative, and therefore measurable, before devising methods for actually measuring them.

Michell (1997, 1999) argues that, since psychological attributes lack additivity, they cannot be continuous quantities of the type encountered in Newtonian physics. According to Michell (1999, p. 71), additivity of an attribute is demonstrated if there is an actual physical process for combining magnitudes of the attribute so that:

1. $w + x = x + w$ for any two magnitudes, w and x, of the attribute, and

2. $(w + x) + y = w + (x + y)$ for any three magnitudes w, x and y of the attribute, and

3. The combined magnitude of the attribute is unchanged if equivalent objects are substituted – objects which have the same magnitudes as the individual objects that are being combined.

For example, combining two objects on a single set of weighing scales would be such a method in the case of weight. Michell's thesis is that there is no such process for combining magnitudes of psychological attributes such as cognitive abilities. As a consequence of his extensive critique of psychological measurement, in which he appeals to the mathematical foundations of measurement, Michell concludes: "These observations confirm that psychology, as a discipline, has its own definition of measurement, a definition quite unlike the traditional concept used in the physical sciences" (Michell, 1997, p. 360). Boring (1929) points out that the pioneers of psychology quickly came to realise that if psychology was not a quantitative discipline which facilitated measurement, psychologists could not adopt the epithet "scientist" for "there would … have been little of the breath of science in the experimental body, for we hardly recognize a subject as scientific if measurement is not one of its tools" (Michell, 1990, p. 7).

The general definition of measurement accepted by most quantitative psychologists and educationalists is that formulated by Stevens (1946), which states: "measurement is the assignment of numerals to objects or events according to rule" (Michell, 1997, p. 360). Stevens realised that, by predicating measurement just upon numerical representation, without the need for additivity, he could ratify and legitimise the measurement practices used by psychologists and liberate them from the need to test the underlying quantitative structure of psychological attributes. One of the consequences of this is that psychologists/educationalists assign numbers to attributes according to some pre-determined rule and do not consider the need to justify the measurement procedures used so long as the rule is followed. For example, psychologists/educationalists use total test scores to measure cognitive abilities without considering the need to justify their rationale for doing so. Stevens' rather vague definition distances measurement practices in psychology and education from measurement in classical Newtonian physics.

A fundamental notion in quantum theory is that a measured attribute can only be understood in the context of the measurement situation. This facet of the quantum realm is in stark contrast to the Newtonian worldview. In Newtonian physics, systems possess their attributes intrinsically and measurement is construed as a process for checking up on some pre-existing property of the system. Niels Bohr, who is widely regarded as the father of quantum theory, identified "unambiguous communication", rather than quantification, as the hallmark of science. To communicate measurement results unambiguously in quantum theory it is essential to factor in a description of the measuring instrument: the entity to be measured and the measuring instrument are non-separable.

The transition in physics from a Newtonian to a quantum theoretical model of reality should have had implications for psychology (Gigerenzer, 1987) and therefore also for education. Bruner (1990, p. xiii) warns that: "The study of the human mind is so difficult, so caught in the dilemma of being both the object and the agent of its own study, that it cannot limit its inquiries to ways of thinking that grew out of yesterday's physics." Bruner is referring to the difficulty of dividing up the mind into that which is the object of introspection (some region of the mind) and the introspecting agent (the mind itself). Bohr also refers to exactly the same problem:

> The epistemological problem under discussion may be characterized briefly as follows: For describing our mental activity, we require, on one hand, an objectively given content to be placed in opposition to a perceiving subject, while, on the other hand, as is already implied in such an assertion, no sharp separation between object and subject can be maintained, since the perceiving subject also belongs to our mental content. (Bohr, 1934/1987, p. 96)

Bohr believed that psychology and quantum theory share a common goal: to use ordinary language, learned through direct experience, to describe what transcends direct experience. He was of the view that he had identified an "epistemological argument common to both fields" (Bohr, 1958/1987, p. 27). Similar claims have been made by other quantum pioneers, for example: "It is the occurrence of similar concepts and thought models in both physics and psychology that makes Pauli so certain that they rest on a foundation of shared structures." (Gieser, 2005, p. 345)

The author draws on aspects of Wittgenstein's later philosophy and Bohr's philosophy of quantum mechanics to argue that education and psychology share the same "measurement problem" as Bohr's so-called "Copenhagen interpretation" of quantum theory. Wittgenstein's philosophy of psychology is utilised since the eminent quantum physicist Werner Heisenberg links the Copenhagen interpretation to Wittgenstein's later philosophy (Stapp, 1972, p. 1114). Wittgenstein's later philosophical writings are used to establish that it is meaningless to ascribe a definite ability to an unmeasured individual; the ascription of a definite ability is only meaningful in a specific measurement context. In particular, if learning is construed as following simple rules to "go beyond the information given" (Bruner, 1996, p. 129), it is evident that Wittgenstein's extensive analysis of rule-following has implications for educational measurement. Accordingly, the following section summarises this important aspect of Wittgenstein's later philosophy. The process of going beyond given information can be illustrated most effectively using mathematical rule-following, but the ideas apply to all rule-following from speaking a language to playing chess.

Wittgenstein's analysis of rule-following

Wittgenstein (2009, §185) considers a scenario in which a child continues the arithmetical sequence 0, 2, 4, 6, … beyond 1000 by writing 1000, 1004, 1008, 1012, and so on. Wittgenstein argues that there is a possibility the child has understood the order to continue the sequence as the community of mathematicians would understand the instruction "Add 2 up to 1000, 4 up to 2000, 6 up to 3000, and so on". Wittgenstein is using this example to communicate that, on the basis of a finite number of illustrations of a rule (the rule for the sequence of even numbers in this case), it is possible to attach multiple interpretations to the rule.

In the latter part of the 20th century, cognitivism superseded behaviourism as the dominant psychological paradigm for understanding mental functions. It therefore seems uncontroversial to view the source of the child's ability to follow a rule as a finite object in his or her mind, and an obvious contender for this in the case of mathematical rule-following is a formula. However, having a formula in mind is insufficient to explain mathematical rule-following behaviour since, in Wittgenstein's view, a rule by itself leads to a paradox in that, on the basis of some interpretation of its requirements, any answer can be made to accord or to conflict with the rule:

> This was our paradox: no course of action could be determined by a rule, because every course of action can be brought into accord with the rule. The answer was: if every course of action can be brought into accord with the rule, then it can also be brought into conflict with it. And so there would be neither accord nor conflict here. (Wittgenstein, 2009, §201)

For example, it is possible for a child to write 10 (as one would expect) or -230 (as one would not expect) when he or she is asked to supply the next term in the sequence 2, 4, 6, 8, ... The child who gives 10 as their answer is attaching the orthodox interpretation $u_n = 2n$, while the child who proffers -230 can claim to be acting in accord with the instruction to continue the sequence by attaching the interpretation $u_n = 2n - 10(n-1)(n-2)(n-3)(n-4)$ since both formulae clearly yield the given first four terms of the sequence.

It is over-simplistic to posit that, if God were to search the contents of the child's mind and uncover the formula $u_n = 2n$, then the child *must* supply 10 as the next term in the sequence 2, 4, 6, 8, ... Analogously, discovery of the formula $u_n = 2n - 10(n-1)(n-2)(n-3)(n-4)$ does not guarantee that the child will proffer -230 as the fifth even number. The child may indeed have the formula $u_n = 2n$ in mind but be unable to use it correctly. Merely having a mental object (the formula in this case) before one's mind may be insufficient since the child may be unable to interpret the object correctly. Perhaps the characteristic which distinguishes the child who can successfully go on to apply a formula correctly from the child who cannot, is the possession of a further mental object which permits the child to attach the correct interpretation to the formula. Unfortunately this will not resolve the conundrum since, if the mind is construed as being populated by mental objects (formulae, images, and so on) which require interpretation, an infinite regress ensues:

> If it [the rule] requires interpretation, that could be done in lots of ways. So how do I tell which interpretation is correct? Does that, for instance, call for a *further* rule – a rule for determining correct interpretation of the original – and if so, why does it not raise the same difficulty again, thereby generating a regress? (Wright, 2001, p. 163)

It is thus evident that an appeal to interpretations does not settle whether the child should offer 10 or -230 as the fifth even number. Having a formula before one's mind does not guarantee the subsequent correct use of that formula, and the introduction of interpretation cannot resolve the problem.

Wittgenstein (2009) anticipates that there may be an attempt to rescue the situation by introducing the idea of a Platonic mechanism in the child's mind, which gives the child access to all potential future applications of a rule without the need for interpretation or, indeed, any act of creativity on the child's part. In this situation, the child would be following a "rules-as-rails" model where the rules "somehow reach ahead of us and determine of themselves every actual and counterfactual proper application" (Wright, 2001, p. 163). Wittgenstein, however, rejected such an explanation of rule-following, as illustrated by his view of mathematical rule-following: "The mathematician is an inventor, not a discoverer" (Wittgenstein, 1978, I, §168).

In Wittgenstein's view, private rule-following within the realm of one's mind is impossible since the public criteria associated with a well-established practice (into which the rule-follower must be enculturated through training) are used to judge correct and incorrect applications of a rule:

> ‘[F]ollowing a rule’ is a practice. And to *think* one is following a rule is not to follow a rule. And that’s why it’s not possible to follow a rule ‘privately’; otherwise, thinking one was following a rule would be the same thing as following it. (Wittgenstein, 2009, §202)

Wittgenstein’s appeal to public, behavioural criteria in his exegesis of rule-following, rather than focusing on associated inner mental states and processes might be construed as indicating that his views are aligned with behaviourism. Wittgenstein, however, denies that he endorses behaviourism when he says “If I speak of a fiction, then it is of a *grammatical* fiction” (Wittgenstein, 2009, §307) in response to the following questions posed by the interlocutor: “Aren’t you ... a behaviourist in disguise? Aren’t you nevertheless basically saying that everything except human behaviour is a fiction?” (Wittgenstein, 2009, §307) Contrary to some readings of Wittgenstein, he does not deny the existence of mental states and processes. Rather, he posits that mental states are inextricably linked to outer behaviour: “An ‘inner process’ stands in need of outward criteria.” (Wittgenstein, 2009, §580) Wittgenstein denies that mental states stand in a direct causal relation to outer behaviour, as per a Newtonian paradigm. Interestingly, he actually steers a path between cognitivism and behaviourism when he asserts, in relation to the nature of mental states, “It’s [a mental state] not a Something, but not a Nothing either! The conclusion was only that a Nothing would render the same service as a Something about which nothing can be said.” (Wittgenstein, 2009, §304) In this quotation, Wittgenstein is rejecting the notion that, in following a rule for example, an individual is guided by some type of mental object (a “something”) which is hidden behind behaviour or indeed that rule-following is reducible without residue to outer behaviour, so that what remains within is a “nothing”.

The following section sets out the rationale for adopting a quantum-theoretical measurement paradigm in education. In addition, Wittgenstein’s later philosophy, and his analysis of rule-following in particular, is used to establish a set of equivalences between the study of educational/psychological predicates and the study of quantum predicates.

The study of educational predicates: quantum analogues

Why should a theory formulated for analysing the behaviour of microentities such as electrons apply to the study of psychological predicates of interest to educationalists such as learning, thinking, understanding, and so on? According to Bohr, the primary concern of the quantum physicist is to communicate unambiguously about aspects of reality which are not visible to the human eye, using ordinary language developed to make sense of the world of everyday experience. The constructs of interest to quantum theorists, such as electrons and photons, manifest themselves in macroscopic measuring devices which can be read by the human eye. Physicists then communicate their experimental findings to colleagues using ordinary language supplemented by some specialist vocabulary.

The predicates of interest to psychologists and educationalists similarly cannot be observed directly but, rather, manifest themselves in the responses of individuals to measuring instruments such as tests. The words “position” and “velocity” function perfectly well when used to characterise the motion of tennis balls, for example, but cannot be used in common-sense ways to describe electrons and photons, which cannot be observed directly. Bohr, however, insists that everyday concepts like position and velocity must be retained in quantum theory but used in a more restrictive way, in the sense that it is necessary to factor in a description of the measuring instrument when referring to a measurement of position or

velocity in the quantum realm. It is argued in the current paper that similar restrictions need to be placed on measurements of abilities in education, i.e. it is necessary to refer to the ability of an individual with respect to a particular question on a test rather than referring to ability as an attribute that exists independently of an act of measurement.

- ***The analogue of quantization***

  In quantum theory, quantized attributes can only take certain discrete values. For example, the polarization of a photon is a quantized attribute since, for each direction, there are only two possibilities: either the photon is completely polarized in that direction or it is completely polarized perpendicular to that direction.

  Wittgenstein rejects the notion that a psychological attribute, such as understanding or ability, is purely a mental process: "In the sense in which there are processes (including mental processes) which are characteristic of understanding, understanding is not a mental process" (Wittgenstein, 2009, §154). This implies that, when one attempts to measure a psychological attribute such as ability, the measurement process cannot simply be a mechanism for checking up on the attribute. Consequently it is meaningless to posit that a general ability, of the type a test purports to measure, can be measured in the Newtonian sense. Consider an *n*-item mathematics test consisting of questions Q_1, Q_2, Q_3, ..., Q_n, where the response to each question is either correct or incorrect. In Wittgenstein's view, "it's not possible to follow a rule 'privately'" (Wittgenstein, 2009, §202), and thus ability *with respect to a given question* must be publicly demonstrated. Wittgenstein advocates the need for an individual to repeatedly demonstrate conformity with a rule-governed practice (such as addition) before one can ascribe the ability to use the rule correctly to the individual. In reality, however, high-stakes educational assessments do not usually test the application of a particular rule on multiple occasions, and it is thus necessary to refer to the ability to apply a rule *in the context of a particular question*, which represents a deviation from Wittgenstein's conception of demonstrating an ability.

  Furthermore, an individual will not have their ability with respect to a given question "in their mind" either before or after they answer the question and, therefore, their ability with respect to the question will be indeterminate both before and after they respond to it. It is noteworthy that the degree of uncertainty pertaining to the ability with respect to a particular question will be greater before the question is attempted than afterwards. The individual's ability with respect to a given question will, however, be maximally certain at the instant they answer it, and this reinforces the fact that it is necessary to refer to their ability *with respect to a particular question*. Therefore, an individual's ability is a relational attribute in the sense that it can only be spoken of in conjunction with a specific question on the test. In conformity with quantum theory, a measurement of ability and the measuring device, i.e., the mathematical practice that is being tested by the question, form an indivisible whole and it is meaningless to refer to the measurement without referring to the measuring device. The need to refer to the ability of an individual with respect to a particular question undermines the rationale for calculating total test scores (which are often purported to represent ability in a particular area), since the summing of item scores on a test is predicated on the notion that each item is measuring the same construct, e.g., "mathematical ability". It is meaningless to combine scores in this way because no single ability exists.

  At the instant when a pupil answers a particular question, Q_r, their ability with respect to Q_r (the measuring instrument) will be known. In other words, the pupil will have

responded either correctly to Q_r (i.e., in conformity with the criteria associated with the relevant practice that is being assessed by Q_r), thus affirming their ability with respect to Q_r, or incorrectly. Since there are only two discrete possibilities for each question, correct or incorrect, the pupil's ability with respect to each test item is a quantized attribute.

- ***The analogue of Heisenberg's uncertainty principle***

Heisenberg's uncertainty principle, which is an important tenet of quantum theory, states that it is impossible to measure precisely both the position and the momentum of a microentity simultaneously.

Consider again an *n*-item mathematics test consisting of questions Q_1, Q_2, Q_3, ..., Q_n, where the response to each question is either correct or incorrect. If an individual responds to question Q_r at time t_r, their ability with respect to Q_r will be maximally certain at time t_r. Again, since Wittgenstein posits that a rule cannot be followed privately (Wittgenstein, 2009, §202), the individual's ability with respect to Q_r will be unknown either before or after t_r. Similarly, if the individual answers question Q_s at time t_s (where $s \neq r$), their ability with respect to Q_s will be maximally certain at time t_s but unknown either before or after t_s. Clearly it is impossible for the individual to respond to two questions simultaneously and, therefore, it is only possible for the ability of the individual to be known with respect to a single question at any instant in time. This is the analogue of Heisenberg's uncertainty principle in quantum theory.

- ***The analogue of wave-function collapse***

In quantum theory, a microentity does not possess its dynamic attributes before a measurement is made. Rather, the microentity exists in a superposition of different states, which correspond to the possible outcomes of the measurement process. A dynamic attribute of the microentity only manifests itself at the point of measurement when the microentity interacts with a measuring device. This interaction gives rise to a sudden and discontinuous jump in the relevant dynamic attribute of the microentity from a superposition of all possible values, with associated probabilities, to one actual measurement result. The probability wave-function, which incorporates information on all of the possible values and their respective probabilities, collapses to yield a single value at the point of measurement.

Wittgenstein offers support for the applicability of the notion of wave-function collapse to the study of psychological and educational predicates when he argues that there is always a "jump" to make between a concept and its application in a particular context:

> In all language there is a bridge between the sign and its application. No one can make this for us; we have to bridge the gap ourselves. No explanation ever saves the jump, because any further explanation will itself need a jump. (Wittgenstein, 1982, p. 67)

Consider once again an *n*-item mathematics test consisting of questions Q_1, Q_2, Q_3, ..., Q_n, where the response to each question is either correct or incorrect. Suppose also that the probability of an individual answering the question Q_r correctly is p_r. Immediately before an individual answers Q_r, their ability with respect to Q_r will, according to Wittgenstein's paradox of interpretation (discussed in the section on rule-following), be characterised by a superposition of two different states corresponding to the two possible outcomes of their response to Q_r:

1. Correct response to Q_r, with associated probability p_r, and

2. Incorrect response to Q_r, with associated probability $1 - p_r$.

According to Wittgenstein (2009), the pupil in isolation could, under some interpretation of the requirements of Q_r, construe either the correct or an incorrect answer to Q_r as correct. For example, Kripke (1982) considers a situation in which an individual who has only previously added whole numbers less than 57 asserts that the correct answer to the mathematical question "What is 68 + 57?" is "5" rather than "125". Kripke (1982) argues that it is possible the individual is using an alternative interpretation of "+" whereby the "+" symbol is being used to denote what he calls the "quus" function instead of ordinary arithmetic addition. Kripke (1982) defines the rule for the "quus" function, which he denotes by "$\oplus$", as follows:

$x \oplus y = x + y$ if $x, y < 57$

$x \oplus y = 5$ otherwise.

Prior to the individual answering the question, they are in a superposition of two states (their answer is both "125" *and* "5"), but when they actually answer the question, the superposition collapses to just one answer (either "125" *or* "5").

Prior to actually answering Q_r, the individual's ability with respect to Q_r is in a superposition of two states simultaneously. However, at the instant when the individual answers Q_r, the superposition of the two states "correct" and "incorrect" will collapse to just one: the actual measurement of the pupil's ability with respect to Q_r. The assessor's knowledge of the pupil's ability with respect to Q_r jumps instantaneously from an uncertain state to a certain state, since the pupil's actual answer to Q_r will be either correct or incorrect. The process of measurement results in a sudden jump in the assessor's knowledge of the pupil's ability with respect to Q_r, from a combination of possible outcomes, each with an associated probability, to one actual outcome. The assessor's knowledge of the relevant ability changes from indeterminate to determinate. The sudden jump in the assessor's knowledge of the pupil's ability with respect to Q_r is the analogue of the process of wave-function collapse in quantum theory.

- ***The analogue of complementarity***

Bohr's principle of complementarity is a central feature of the Copenhagen interpretation of quantum theory. In the quantum world some observations can never be made simultaneously. For example, one cannot observe a microentity as both a particle and as a wave at the same time. A microentity behaves as a particle when it is observed, but is characterised by a probability wave between observations. The two situations cannot occur simultaneously, i.e., they are mutually exclusive, but both are necessary to fully describe the microentity's nature. Bohr believed that the complementarity principle accommodates the paradoxical nature of the wave-particle duality exhibited by microentities. He contended that "evidence obtained under different conditions cannot be comprehended within a single picture, but must be regarded as *complementary* in the sense that only the totality of the phenomena exhausts the possible information about the objects" (cited in Kumar, 2009, p. 242).

The complementarity principle states that the separate descriptions of the quantum realm obtained from different measurement events, or measured versus unmeasured systems,

cannot be combined into a single comprehensive description of reality. Rather, it is only possible to give a sequence of complementary descriptions which correspond to different measurement situations, or to the measured versus the unmeasured system. Honner (1987, p. 59) refers to the type of complementarity that exists between the descriptions of a system corresponding to different measurement situations as "weak" complementarity. For example, if a photon passes through a sequence of n polarizing disks then, to completely describe the photon's polarization with respect to each of the disks, it is necessary to give n distinct statements, all of which are complementary (rather than contradictory) to each other. Honner (1987, p. 59) contends, however, that the complementarity that exists between the descriptions of a measured versus unmeasured system (e.g., the wave and particle manifestations of an electron corresponding respectively to the unmeasured and measured states of the electron), which he refers to as "strong" complementarity, is more closely aligned with Bohr's original conception of complementarity. Accordingly, the analogue of strong complementarity in the study of psychological and educational predicates will be discussed in this section.

Wittgenstein (2009) demonstrated that there is a radical asymmetry between first-person present tense use and third-person present tense use of psychological concepts. First-person use and third-person use of mental predicates are asymmetrical with respect to their methods of verification since third-person use is based upon behavioural criteria, whereas first-person use is not. However, despite this asymmetry in their methods of verification, first-person and third-person use of mental concepts have, according to Wittgenstein, a univocity of meaning since they are linked to common behavioural criteria.

In Wittgenstein's view, asymmetry of use but univocity of meaning is a feature of all psychological predicates including intentional predicates such as thinking, learning, understanding, and so on. For example, it is meaningful to speak of correct or incorrect in third-person use of the word "ability", but it is meaningless to speak of correct or incorrect in first-person use of the word "ability" (as there are no criteria to adjudicate on first-person use). Similarly third-person use of the word "ability" is based on description, whereas first-person use is not, so that first and third-person uses of an intentional predicate are mutually exclusive. However, both first-person and third-person perspectives are necessary to give a full account of an intentional predicate. As Bohr noted, "debates between behavioristic and introspectivist positions in psychology can be resolved by recognizing that both are necessary to complement each other for a full account of human experience" (Faye and Folse in Bohr, 1998, p. 19). First-person use of an intentional predicate precludes third-person use and the two uses are therefore not contradictory but, rather, complementary. Therefore, the asymmetry of first-person and third-person uses of intentional predicates, such as learning, is analogous to strong complementarity in quantum theory.

- ***The analogue of irreducible uncertainty and objective probability***

Quantum theory cannot determine the exact behaviour of microentities but, rather, only permits probabilities of different possible outcomes to be calculated. Quantum uncertainty is irreducible in the sense that it cannot be reduced by acquiring further information about the system being considered, and the probabilities associated with the uncertainty are objective rather than subjective. The resort to probabilities is not associated with ignorance, and the irreducible uncertainty is simply a fact of the natural world. No theoretical approach has hitherto managed to circumvent the irreducible uncertainty in the quantum realm.

However, the situation in classical physics is very different from the quantum world. In classical Newtonian mechanics, it is only necessary to resort to probability when there is incomplete information regarding a particular system. Consider, for example, the situation that arises when an unbiased coin is tossed. If all of the parameters pertaining to the coin's motion were known, e.g., initial position, initial velocity, coefficient of restitution between the coin and the surface upon which it lands, etc. were known, it would be possible to predict, with certainty, the outcome of the coin-tossing experiment (Strzałko *et al*., 2008). In practice, however, at least some of these parameters will be unknown, and it is this "classical ignorance" which leads to probabilities being invoked to describe the experimental outcomes. Probabilities of this type are subjective since the uncertainty with which they are associated would be reduced if further information were available. Indeed, with complete information regarding all of the parameters influencing the motion of the coin, the uncertainty would yield to certainty.

The Cartesian conception of the mind leads to the belief that, in the study of educational predicates such as thinking, the probabilities an individual associates with another person's thoughts are subjective, since the individual does not have direct access to the private mental states of the other person. Those aligned with the Cartesian view posit that there would be no uncertainty, and therefore no need to resort to probabilities, if it were possible for the individual to have direct access to the mental states of the other person.

In Wittgenstein's rejection of Cartesianism, one's uncertainty in describing another person's mental states does not flow from ignorance of those mental states. According to Wittgenstein, when a person expresses a thought, for example, they are not describing an inner state with which the expression can be checked for accuracy. Wittgenstein's private language argument rejects the existence of

> … a language in which a person could write down or give voice to his inner experiences – his feelings, moods, and so on ... The words of this language are to refer to what only the speaker can know – to his immediate private sensations. So another person cannot understand the language. (Wittgenstein, 2009, §243)

This suggests that the mental states of a person are not directly comparable with the person's subsequent actions and, therefore, that the uncertainty pertaining to the person's mental predicates cannot be reduced by inspecting their mental states.

Furthermore, if thoughts were self-contained, isolated brain processes that could be analysed using a brain imaging device, this would violate the characteristic first-person/third-person asymmetry which Wittgenstein (2009) identified as a feature of all mental predicates. First-person/third-person asymmetry refers to the fact that third-person ascriptions of psychological predicates are based upon the use of criteria whereas first-person ascriptions are not. If the asymmetry is the defining property of the mental, then mental phenomena cannot be captured in brain processes:

> But the view that thoughts and feelings are brain-processes *abolishes* this logical difference. If this view were true, you and I would stand on the *same* level in regard to what I think and feel. In order to ascertain my thoughts and feelings you *and* I would equally have to rely on advanced technology and scientific theory. (Malcolm, 1986, p. 191)

Alas, physical entities such as brain processes are governed by first-person/third person *symmetry* rather than asymmetry since both first-person and third-person ascriptions of

physical predicates entail the invocation of criteria. “After all, we say that Jones is six foot tall for the very same reason we say that we are.” (Suter, 1989, p. 153) Entities characterised by first-person/third-person asymmetry cannot be construed as identical to entities characterised by first-person/third-person symmetry.

Therefore, the probabilities associated with mental predicates are objective, rather than subjective, since the uncertainty does not arise as a consequence of either human or instrumental limitations. Even if direct access to inner mental states were possible, it would reveal nothing and, consequently, the uncertainty associated with mental predicates is irreducible in the same way as the uncertainty in the quantum realm.

- ***The analogue of local hidden variables***

Einstein was an objective realist who believed strongly in a deterministic universe in which it is always possible to trace a clear path from local cause to effect. He refused to accept that quantum theory represents a complete fundamental theory of the natural world. A number of quantum theorists hold that Einstein insisted there must be “elements of reality” which, if incorporated into quantum theory, would make it a more complete theory of physical reality. Einstein’s “elements of reality” are construed by physicists to refer to local hidden variables, i.e., Platonic “look-up lists” which dictate the values of the dynamic attributes of a microentity in every possible measurement scenario. For example, according to Einstein, an unmeasured photon possesses a type of instruction list that identifies the photon’s polarization along every possible angle before it is actually measured. However, Bell’s theorem (Bell, 1964) proves conclusively that local hidden variables of this type do not exist.

Consider again an *n*-item mathematics test consisting of questions Q_1, Q_2, Q_3, ..., Q_n, where the response to each question is either correct or incorrect, and suppose that an individual is about to answer a particular question Q_r on the test. If the equivalent of Einstein’s local hidden variables were to exist in the context of educational measurement, then the individual’s ability with respect to Q_r would be definite *in advance of* their answer to the question. In other words, the individual would have a hidden, internal “look-up list” in their mind which would guide their response to Q_r. According to this view, whilst the assessor, due to their ignorance of these hidden variables, does not know how the individual will respond to Q_r before it is actually answered, the individual’s response could be predicted if this ignorance were eradicated. However, Wittgenstein’s later philosophy suggests that such a view is implausible.

Wittgenstein counsels against looking for hidden causes of the behaviour associated with mental phenomena since he claims such a quest leads to conceptual confusion:

> Now we try to get hold of the mental process of understanding, which seems to be hidden behind those coarser, and therefore more readily visible, concomitant phenomena. But it doesn’t work; or, more correctly, it does not get as far as a real attempt. For even supposing I had found something that happened in all those cases of understanding, why should that be the understanding? ... And if I say it is hidden – then how do I know what I have to look for? I am in a muddle. (Wittgenstein, 2009, §153)

In particular, he denies that the mental life of an individual is guided by latent, internal representations such as mental images. Wittgenstein concedes that mental images may be accompaniments of thinking but he vehemently denies that such images actually *determine*

how thoughts are expressed. If mental images were to determine the thinking of an individual, then Wittgenstein's paradox of interpretation (discussed in the section on rule-following) would lead to the bizarre consequence that the individual in isolation might always be mistaken about their thoughts.

In Wittgenstein's (2009) view, a rule-follower is neither following a rule arbitrarily nor being guided by local hidden variables such as mental images. Rather, the rule-follower's guidance is non-local: their past training. The training is non-local since it is not stored inside the rule-follower as some form of mental object, and hidden in the sense that the training itself (rather than its consequences) does not have any physical manifestation that is detectable by physically examining the rule-follower. Wittgenstein acknowledges that human beings have a repertoire of natural responses that can be moulded through training. However, he rejects the notion that training simply induces physical changes to neurological pathways in the brain that could be measured by an appropriate imaging device, since this would undermine the first-person/third-person asymmetry that is characteristic of psychological predicates (see above). Whilst changes to brain states obviously could be measured, the entity represented by the measurement would be governed by first-person/third-person symmetry and, therefore, it would be at variance with our common conception of psychological attributes. If one attempted to measure, for example, understanding using such advanced technology, the measurement would actually be of an entirely different predicate, understanding*, say, that is incongruous with our everyday conception of understanding. Furthermore, as described in the section on rule-following, Wittgenstein rejected the idea of a Platonic mechanism in an individual's mind that somehow gives access to all future applications of a rule.

Therefore, an individual who is just about to answer question Q_r on the *n*-item mathematics test cannot be guided by some*thing* in mind; the individual does not have a hidden "look-up list" that predetermines their answer to Q_r. Wittgenstein's philosophical discussions on the nature of meaning and understanding lead to the same conclusion for psychology and education as Bell's theorem led to for quantum theory: local hidden variables that guide behaviour are untenable.

The structural parallels that have been established between the study of educational/psychological predicates and quantum theory suggest that a quantum theoretical framework would provide a more secure philosophical basis for educational measurement than the Newtonianism that underpins contemporary measurement models. The essential features of this alternative measurement paradigm for educational predicates are summarised in the following section.

Quantum measurement paradigm for educational predicates

Traditionally educationalists have viewed measurement through a Newtonian lens, in the sense that measurement is construed as a process for checking up on an innate psychological attribute of a person: their ability. In the Newtonian tradition, ability is viewed as an intrinsic attribute of an individual which stands in a causal relation to their responses to the questions on a mathematics test, for example. The responses are judged to be "correct" or "incorrect" according to an independent standard that is external to the individual.

Consider again an *n*-item mathematics test consisting of questions Q_1, Q_2, Q_3, ..., Q_n, where the response to each question is either "correct" or "incorrect". In the quantum theoretical

framework for educational measurement, it is meaningless to speak of ability as an intrinsic property of an individual. Rather, it is only meaningful to refer to the individual's ability with respect to a given question at the instant when they actually respond to it. Therefore, an individual's responses to the *n* questions will constitute *n* discrete measurements of their mathematical ability with respect to each question.

Prior to measuring an individual's ability with respect to a given question, they are in a superposition of two states: their answer is both "correct" *and* "incorrect". The superposition is ontological rather than epistemological. It is not that one does not know which state the individual is in, but they really *are* in both states simultaneously. When a measurement is made, the superposition collapses to yield one actual measurement result: the answer is either "correct" *or* "incorrect", and the uncertainty pertaining to which outcome will be actualised in the measurement process is irreducible. During the measurement, the individual and the measuring device, i.e. the practice assessed by the question, form a unified and non-separable system that actually influences the measured value of the ability.

If an individual answers all *n* items on the hypothetical mathematics test, their ability will be definite on *n* distinct occasions (when they are answering each of the *n* items) and indeterminate between consecutive items (when they are not responding to a specific item). According to Bohr's principle of complementarity, it is necessary to give *n* distinct descriptions (one corresponding to the individual's ability with respect to each test item) to completely capture the individual's mathematical ability. It is impossible to combine the *n* statements since they are all necessary to give a complete description of the individual's mathematical ability. A consequence of this re-construal of educational measurement is that total test scores are meaningless since the rationale for summing item scores on a test is that each item is measuring the same construct, e.g., "mathematical ability". The scores cannot be combined in the quantum measurement paradigm because no single ability exists.

Conclusion: Implications for validity and the future of educational measurement

Educationalists consider validity to be a critical issue in assuring the accuracy of the instruments which they use to measure ability levels. For example, in 1989 the British Psychological Society's Steering Committee on Test Standards defined validity as "the extent to which a test measures what it claims to be measuring, the extent to which it is possible to make appropriate inferences from the test score" (cited in Coaley, 2009, pp. 129-130).

Educationalists contend that a test is valid if variations in the ability measured by the test are causally related to variations in the measurements derived from the test. The central tenet of validity in the context of educational measurement is the facility to abstract a measurement of ability away from the measuring instrument. The concept of validity evidently presupposes that an individual possesses an intrinsic ability level which is independent of the instrument used to measure it.

The quantum measurement paradigm presented in the current paper stresses that an individual's ability is actually a joint property of the individual and a particular test item. It is simply meaningless to divorce the ability from the measuring instrument. The author thus contends that the concept of validity as it is traditionally conceived in educational measurement is a misnomer.

The erosion of the notion of validity in educational measurement has grave implications for the legitimacy of the high-stakes assessments that currently dominate the educational landscape. Assessment results are, for example, used to award qualifications that determine future educational or vocational pathways of students. The results obtained by students in assessments are also used to gauge individual teacher quality, to hold schools to account for the standards achieved by their students, and to compare international education systems. Whilst it is acknowledged that transition to a quantum theoretical framework would not lead to the demise of educational assessment, the author believes that the focus on high-stakes assessments should be reduced. Ideally such assessments should be reformed to become as 'low-stakes' as possible through, for example, placing greater emphasis on qualitatively describing what students can do in particular contexts, rather than on attempting to quantify their educational achievements.

High-stakes assessments may, however, be justifiable in some contexts, despite the fact they do not yield valid measures of ability. There may still be strong links between a measurement produced by a test and a future expected outcome, which are useful for practical inductive reasoning, such as for making employment decisions, even if the test does not provide an accurate, ongoing, present-tense descriptor for the entity it purports to measure. It is the predictive strength of test scores that should determine if high-stakes assessments are justifiable in particular contexts. The ultimate consequence of the measurement model described in the current paper is to undermine some of the pro high-stakes assessment rhetoric that afflicts education, such as when authorities claim with an unjustifiably high degree of confidence what assessment data do or do not show, or they make epistemic claims about ability levels on the basis of test scores.

References

Bell, J.S. (1964). On the Einstein Podolsky Rosen paradox. *Physics*, *1*, 195-200.

Bohr, N. (1934/1987). *The philosophical writings of Niels Bohr volume 1: Atomic theory and the description of nature.* Woodridge, CT: Ox Bow Press.

Bohr, N. (1958/1987). *The philosophical writings of Niels Bohr volume 2: Essays 1933-1957 on atomic physics and human knowledge.* Woodridge, CT: Ox Bow Press.

Bohr, N. (1998). *The philosophical writings of Niels Bohr volume 4: Causality and complementarity* (Supplementary papers edited by J. Faye & H.J. Folse). Woodridge, CT: Ox Bow Press.

Boring, E.G. (1929). *A history of experimental psychology.* New York, NY: Century.

Bruner, J.S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Bruner, J. (1996). *The culture of education*. Cambridge, MA: Harvard University Press

Coaley, K. (2009). *An introduction to psychological assessment and psychometrics*. London, England: Sage Publications.

Gieser, S. (2005). *The innermost kernel*. Berlin, Germany: Springer-Verlag.

Gigerenzer, G. (1987). Probabilistic thinking and the fight against subjectivity. In L. Krüger, G. Gigerenzer, & M.S. Morgan (Eds.), *The probabilistic revolution – Volume 2: Ideas*

in the sciences (pp. 11-33). Cambridge, MA: The Massachusetts Institute of Technology Press.

Honner, J. (1987). *The description of nature: Niels Bohr and the philosophy of quantum physics*. Oxford, England: Oxford University Press.

Kripke, S.A. (1982). *Wittgenstein on rules and private language*. Oxford, England: Blackwell.

Kumar, M. (2009). *Quantum: Einstein, Bohr and the great debate about the nature of reality.* London, England: Icon Books.

Lagemann, E.C. (2000). *An elusive science: The troubling history of education research*. Chicago, IL: University of Chicago Press.

Malcolm, N. (1986). *Wittgenstein: Nothing is hidden*. Oxford, England: Blackwell.

Michell, J. (1990). *An introduction to the logic of psychological measurement.* Mahwah, NJ: Lawrence Erlbaum.

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology, 88*, 355-383.

Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept.* Cambridge, England: Cambridge University Press.

Stapp, H.P. (1972). The Copenhagen interpretation. *American Journal of Physics*, *40*, 1098-1116.

Stevens, S.S. (1946). On the theory of scales of measurement. *Science*, *103*, 677-680.

Strzałko, J., Grabski, J., Stefański, A., Perlikowski, P., & Kapitaniak, T. (2008). Dynamics of coin tossing is predictable. *Physics Reports*, *469*, 59-92.

Suter, R. (1989). *Interpreting Wittgenstein: A cloud of philosophy, a drop of grammar*. Philadelphia, PA: Temple University Press.

Wittgenstein, L. (1978). *Remarks on the foundations of mathematics* (Edited by G.H. von Wright, R. Rhees & G.E.M. Anscombe; translated by G.E.M. Anscombe; 3rd edition). Oxford, England: Blackwell.

Wittgenstein, L. (1982). *Wittgenstein's lectures, Cambridge, 1930-1932: From the notes of John King and Desmond Lee* (Edited by D. Lee). Oxford, England: Blackwell.

Wittgenstein, L. (2009). *Philosophical investigations: The German text with an English translation* (Edited by P.M.S. Hacker & J. Schulte; translated by G.E.M. Anscombe, P.M.S. Hacker & J. Schulte; 4th edition). Chichester, England: Wiley-Blackwell.

Wright, C. (2001). *Rails to infinity*. Cambridge, MA: Harvard University Press.