

ASSESSMENT WITHOUT LEVELS

INVESTIGATING THE IMPACT OF RECENT CHANGES IN THE NATIONAL CURRICULUM ON PHYSICAL EDUCATION TEACHERS' PRACTICE

Liam McGowan

INTRODUCTION

Assessment within physical education (PE) should support meeting the physical, social, cognitive and psychological needs of students – a complex task (Tinning and Rossi, 2013; Hanson, Mullins and Modi, 2017). Understanding current assessment practice and its effectiveness within PE will afford opportunities for improved decision-making from policymakers and schools to improve outcomes for students. In September 2014, the statutory assessment levels embedded in schools' practice were abolished by the government. With no published academic material discussing the impact of this policy change on PE in England, this investigation aims to identify and discuss the impact of that policy change and subsequent teachers' practice with students aged 11-14. The identified age group is important as there are no national academic qualifications for this age group.

METHOD

Through an interpretivist paradigm, a qualitative method adopting semi-structured interviews (recorded via Skype) followed by an inductive, thematic analysis was used for this investigation (Braun and Clarke, 2006). This allowed themes to be established and linked to the data (Patton, 1990). Visual representations were created to establish links between the data and themes.

Bourdieu's concept of habitus (deeply ingrained habits, skills and dispositions) was considered during analysis. This approach allowed for the themes to be explored while considering the complexity of school environments and systems. Habitus was also adopted as a concept to structure analysis, considering personal experiences and participant interactions, group interactions and organisational interactions, and systemic interactions.

There was a purposive mixed geographical sampling of participants. To provide valid reflection of current practice, all participants (n=8) had to meet the eligibility criteria.

FINDINGS AND DISCUSSION

Participants were forthcoming in sharing personal views and feelings on assessment practices. Themes included: student understanding; views on school

assessment; views on changes; summative assessment system; decision-making; school requirements; and influencing factors. Although influencing factors and decision-making were repeating themes, the overarching discourse shone a light on the systems teachers and their schools had created, and the effectiveness of the practices they chose to engage in.

The penultimate and final theme maps are shown in Figure 1a and Figure 1b.

Figure 1a

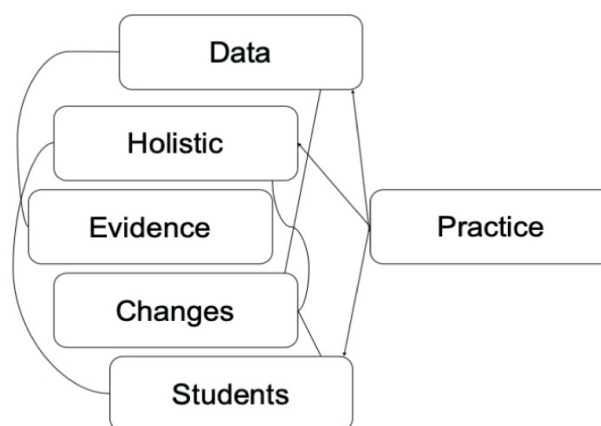
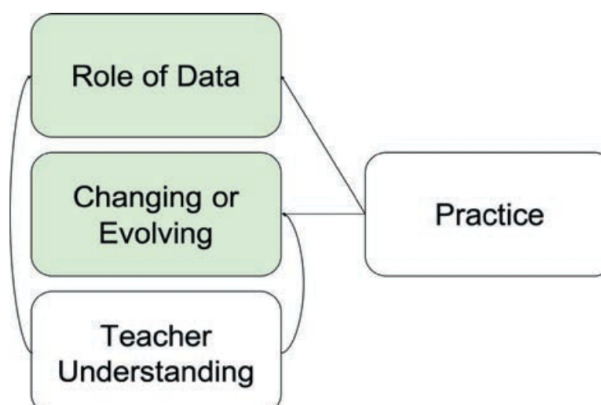


Figure 1b



The data analysis resulted in the identification of two key themes: the role of data and changing or evolving?

ROLE OF DATA

This theme is defined as the influence of assessment procedures in school. All participants engaged in the process of collecting summative data on student performance. Seven participants collected at least some data that was quantified and submitted to more senior personnel within their school. These interactions with summative assessment were varied.

“Every term we do a data drop. The data it will be from one to nine (GCSE type grading).”
(P7, paragraph 69)

While this highlights a teacher was unfazed by the use of GCSE grading with an incorrect age group, other teachers expressed concerns about their interactions with summative data collection.

“The school use thresholds. So they call them learning thresholds. Now I think it's still levels. The school use foundation, developing, secure and then excellence... So the thing with that for me is, each of those thresholds is given a value in numerical value. And this is purely for SLT for data etc. So, for example, foundation is worth 1, developing is worth 2, secure is 3, excellence is 4.”
(P6, paragraph 12)

The removal of levels was supposed to eliminate this poor practice. Worryingly, though, poor practice was morphing into inconsistent and increasingly ineffective practices:

“Data that goes to SLT and parents is one holistic number, which is actually broken into subdivisions, so 1a 1b 1c.”
(P5, paragraph 58)

The morphing also allowed practice to be disguised in qualitative appearance:

“It is very much the language attempting, developing, refining and skillful, and then it is the seven key skills which we use the most.”
(P1, paragraph 104)

To the observer, it may appear that, in attempting to judge attainment and measure progress, teachers are engaged

with helpful and professional processes. However, it was only P8 who made significant alterations to assessment practice and procedures. These changes had begun even prior to the removal of levels, but their removal made establishing the desired evidence-based practice possible.

“We don't have to do the summative judgements anymore. We still write written reports, so we write about their progress in learning domains and where in areas they can still further develop to help become physically educated and leading a physically active lifestyle. We do that twice a year, and we tend to do that, and before we do that we do a written self-reflection from the boys within the four domains, so where they feel they have made progress and where they still feel they need to make progress, and then teacher judgement on that. So we have the self-reflections and the data taken from the Year 7, 9 and 11 questionnaires, and we try and triangulate that with voluntary interviews from the boys.”
(P8, paragraph 88 and 92)

Therefore, despite the Department for Education (DfE) and William (2015) recognising that making ‘best fit’ judgements was an inappropriate assessment practice, all but one of the teachers confirmed this practice still exists in a new guise.

These procedures exemplify the power of existing professional interactions in limiting the development of effective practice. This raises a poignant question: if teachers are engaged in data collection, why are they not more critical of this? Furthermore, what is resulting in teachers’ assessment literacy levels and their ability to critique their professional habits effectively?

Analysing organisational interactions regarding how these practices were formed, it is unsurprising that accountability – in this instance, external exams – drew the significant attention of practitioners.

“We've spent a lot of time looking at our assessment and marking and doing CPD GCSE, BTEC, those kind of areas but Key Stage 3 I would say we haven't as a department or a school been given a lot of guidance.”
(P3, paragraph 84)

Considering the pressure of accountability, programmes of study with linear assessment models that were adapted to accommodate the desire to measure progress between students’ entry and exit of secondary school became prevalent.

“We have a programme of study tracker, so they get levelled [as] emerging, developing, secured and mastered. So, we always have two places that we put in the levels which is quite frustrating. But that is changing I believe to I think they're trying to get rid of the levels, the actual numbers, and move to a below, on target, and above. I think because our flight path at our school, when they get their progress reviews, end up getting quite a lot of red because our progress reviews are colour-coded.”
(P4, paragraph 12)

This highlights problems with a poorly designed assessment systems and that schools are even willing to change the system to fit poor practices. Even when some departments spent time designing assessment systems, they did not necessarily yield results or processes that differed from the National Curriculum with levels.

“We looked at what a load of other schools are doing, up and down the country and again abroad, and what we found was regular kind of data drops... were really important.”
(P2, paragraph 8)

It is almost ironic that, despite awareness of skill acquisition and progress occurring in a non-linear fashion (Chow *et al.*, 2013), existing assessment practice ignores the maturation process many students experience in Key Stage 3 (Lloyd and Oliver, 2012) and expect linear progress to occur. Only P8, who did not report summative data to more senior school personnel, consciously led their department with an evidence-based approach to assessment in PE and escaped this ‘data drop’ practice.

“We have tried to make them evidence-based. When I say evidence-based I mean based on good theory, based on good research, based on objective data. So our physical competency assessment. Based on subjective data, i.e. the children's self-reflections... the combined experience of the rest of the PE department and also involving [removed name] University, who are our local ITT provider, and for the last two or three years academics from [removed name] University.”
(P8, paragraph 132)

Therefore, questions are raised as to whether a competent, critical approach is being adopted to the sources of evidence and support when designing summative assessment practices.

At the wider system level, the removal of levels attempted to consider and understand the complex interactions occurring in schools and reduce rather than add to teacher workload (Christodoulou, 2016). However, lack of guidance, coupled with opportunistic entrepreneurs – based in and out of academy trusts – pushing new assessment systems, appears to be fuelling and compounding problems in guiding teachers toward ineffective practice rather than leading to the emergence of positive outcomes.

“ Our Director of Curriculum actually commented to me the other week actually that our [name of assessment model] is one of the best he’s ever seen and we will be selling it to other academies in our trust. ”
(P4, paragraph 76)

The national assessment system that existed between 1995 and 2014 meant the education workforce behaved in a certain way. Altering policy and expecting improved change without any intervention was, at best, optimistic. Therefore, the role of data and its influence on practice is potentially distracting from more effective formative and summative practices. Is the data being captured accurate and resulting in positive adaptations to practice that are beneficial to the learner or is it detracting from potentially increased time for curating effective learning opportunities?

CHANGING OR EVOLVING?

This theme is defined as the effectiveness of assessment practices in school. In 2014, national policymakers accepted change would result in perturbations in practice and school policy (Christodoulou, 2016). However, findings demonstrate increased variance in practice which has led to increasingly unknown levels of effectiveness.

The teachers who were interviewed felt positive about the national policy change. It created the opportunity for an increased focus on alternative learning domains rather than assessment focusing purely on physical performances. This focus prevailed despite the sparse guidance on what students in Key Stage 3 should be taught from a holistic viewpoint. Exploring holistic approaches to assessment allowed teachers to consider the purpose of their PE departments, with some choosing to focus on explicit teaching and/or assessment of a range of non-physical aspects, such as leadership, motivation, evaluation, thinking skills, social skills, creativity and commitment.

One participant’s school assessed the four learning domains of affective, cognitive, psychomotor and social. Two participants’ schools created rubrics based on an idea promoted by the Youth Sport Trust which included Thinking Me, Social Me, Physical Me, Personal Me and Healthy Me. One participant’s school used an idea promoted by the Association for Physical Education:

Head, Heart and Hands. These ideas are informed by literature regarding domains of learning, and support an alternative, holistic approach to PE, encapsulating the nineteenth century John Ruskin quote adopted by John Haden Badley’s 1893 Bedales School vision of educating head, hand and heart (Bedales School, 2019).

The practice of both students and teachers assessing student competency and learning in multiple domains, usually on individual subjective opinion, raises queries regarding the validity and quality of feedback and assessment judgements occurring and to what extent they support learning (Oliver and Kirk, 2014).

“ I print out their faces and I do the kind of Me in PE or the learning objectives variety of different outcomes on the board or on the gym wall, and they stick their face. We call it place your face. It shows your progress. ”
(P2, paragraph 28)

Some teachers had intentionally embedded evidence-informed practice within their assessment, pedagogy and curriculum design.

“ The way it (Essential Movement Programme) is designed means they assess constantly, having to actually support their peers’ movement on three criteria, technique, control and range of movement, and that is the language they need and how they should observe. ”
(P8, paragraph 80)

Notably, this level of detail around informed practice was disclosed by one participant, and it was in the only school that did not have to report summative assessment judgements to senior personnel. In this and one other school, further change was evident as they implemented a models-based approach, whereas other schools continued with various traditional approaches, teaching physical activities in four to eight-week blocks, which has known problems (Kirk, 2011).

Despite evidence of emerging positive practice, in some schools where more senior personnel dictate policy, problems in practice began to unfold. With freedom to adapt, and with external pressures of accountability mounting, the schools in the study graded as ‘requires improvement’ by Ofsted found themselves with less freedom to explore developing practice at Key Stage 3. This reflects previous research published



by Ofsted where 85 per cent of interviewed senior leaders said they gave less priority to Key Stage 3 compared to Key Stages 4 and 5 (Ofsted, 2015).

Therefore, with examinations dictating schools' habits, it is unsurprising to see practice suitable for examination-based subjects inappropriately transferred to or applied in the context of Key Stage 3 PE.

“We used the AQA (exam board) GCSE criteria, and obviously some pupils won't access that, but we reworded and made a simplified version of what they were looking for.”
(P5, paragraph 110)

This again revealed teachers' and schools' inability to progress beyond current practices and their associated problems.

At a systemic level, policymakers must acknowledge the rise and influence of teaching schools alliances and academy trusts on professional learning habits. With a reduction in local authority support, following the removal of nearly all advisors, and a clear direction to academies that they should seek to generate extra funding themselves, senior leaders of these institutions can influence at a local level with limited knowledge of effective practice or practice that lacks robust empirical evidence. Therefore, schools must be wary of any guidance that relies only on local, case-based evidence as they steer their institutions in directions that may be unwise. Attempts to measure, recreate and adapt systems that quantify attainment and progress in these areas recreate the quantitative and subsequent organisational inefficiencies and problems that previous practice – with levels – was creating (Ofsted, 2018).

Evidently PE departments are increasingly adopting holistic approaches to assessment. This has not been recognised by policymakers and school inspectors. Without understanding at a system-wide level, creating an environment that is conducive to effective professional growth becomes increasingly challenging. Within this environment, school leaders, policymakers and politicians must be aware of the consequences of their actions, such as the impact of academisation and marketisation (Sperka and Enright, 2017) on current assessment practices, and potential antithetical outcomes.

Seemingly the ecological demands at all levels of the education system result in difficult conditions for PE practitioners to develop effective practice. Senior leaders

must consciously engage in practice to evolve and escape old habits by loosening the leash on PE departments adhering to whole school assessment policies while developing organisational and departmental understanding of assessment literacy.

CONCLUSION

Adaptations in teachers' assessment practice as a result of the removal of statutory assessment levels have been dictated by the environment and socialisation within these environments. Collecting assessment data still has a strong presence in teachers' practice. More positively, a focus on developing increasingly holistic approaches to the curriculum has occurred. This has not been recognised by state institutions, namely the DfE and Ofsted. There is, however, little evidence that holistic approaches to assessment, like those being implemented by some of the study's participants, yield improved outcomes for the students.

The recommendation from this study is that, with external support – ideally including academic personnel, PE departments should seek to commit time to consult internal and external sources, including academic literature, to improve departmental assessment literacy, assessment practice and subsequent outcomes for students (Borghouts *et al*, 2019). This practice should accommodate the understanding of non-linear student learning, while engaging students in continual assessment processes across integrated learning domains. It is also important to be explicit in recommending that the use of grading based on uninformed assessment criteria for any learning domain should not be used, particularly when judgements result in overall summative or 'best fit' as this would not be valuable (Kime, 2017). Schools, researchers and state institutions should investigate the effectiveness of the holistic assessment practices being implemented before encouraging others to mirror or apply practices in their own school contexts. Further research into existing assessment practices in specific types of schools is required to yield further understanding from which effective decisions can be made about future interactions at a classroom, organisational and system level.

Please note that this article is an abridged version of the study. In particular, the section on 'methods' was cut by the editor in the interests of space. Study limitations include the lead researcher being a practising PE teacher and, therefore, not separate or independent from the field of study.

ACKNOWLEDGMENTS

Ethics for this study were granted by the University of Birmingham's School Ethics Committee on 5th December 2018. I am grateful to all of the participants for freely volunteering their time and considered responses during their interviews. ■

REFERENCES

- Bedales School. 2019. History. Bedales.org.uk. Available at: <https://www.bedales.org.uk/our-schools/history>
- Black, P. and Wiliam, D., 2003. 'In praise of educational research': formative assessment. *British Educational Research Journal*, 29(5), pp.623-637.
- Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), pp.77-101.
- Christodoulou, D., 2016. *Making Good Progress? The future of Assessment for Learning*. Oxford: Oxford University Press.
- Clarke, V. and Braun, V., 2016. Thematic analysis. *The Journal of Positive Psychology*, 12(3), pp.297-298.
- DfE 2015. Commission on Assessment Without Levels: final report. GOV.UK. Available at: <https://www.gov.uk/government/publications/commission-on-assessment-without-levels-final-report>
- Hanson, M., Mullins, E. and Modi, N., 2017. Time for the UK to commit to tackling child obesity. *BMJ*, p.i762.
- Kime, S. *et al.*, 2017. *What Makes Great Assessment?* Durham, Evidence Based Education.
- Kirk, D., 2011. *Physical Education Futures*. 1st ed. London: Routledge.
- Kvale S., 1996. *InterViews: An Introduction to Qualitative Research Interviews*, Sage Publications, California
- Lloyd, R. and Oliver, J., 2012. The Youth Physical Development Model. *Strength and Conditioning Journal*, 34(3), pp.61-72.
- Lo Iacono, V., Symonds, P. and Brown, D., 2016. Skype as a Tool for Qualitative Research Interviews. *Sociological Research Online*, 21(2), pp.1-15.
- Ofsted 2015. Key Stage 3: the wasted years?. GOV.UK. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/459830/Key_Stage_3_the_wasted_years.pdf
- Ofsted 2018. Curriculum research: assessing intent, implementation and impact. GOV.UK. Available at: <https://www.gov.uk/government/publications/curriculum-research-assessing-intent-implementation-and-impact>
- Oliver, K. and Kirk, D., 2014. Towards an activist approach to research and advocacy for girls and physical education. *Physical Education and Sport Pedagogy*, 21(3), pp.313-327.
- Patton, M.Q., 1990: *Qualitative evaluation and research methods*, second edition. Sage.
- Power, E., 1999. An Introduction to Pierre Bourdieu's Key Theoretical Concepts. *Journal for the Study of Food and Society*, 3(1), pp.48-52.
- Reay, D., 2004. 'It's all becoming a habitus': beyond the habitual use of habitus in educational research. *British Journal of Sociology of Education*, 25(4), pp.431-444.
- Tinning, R. and Rossi, A., 2013. Thinking about complexity thinking for physical education. In: A. Owens, T. Hopper and J. Butler, ed., 1st ed. London: Routledge, pp.195-206.

Liam McGowan is currently Head of PE at Priory School, Edgbaston, working with pupils aged 3-18 years old. He is also a former afPE Board Member. This article is a summary of a submitted Masters dissertation at the University of Birmingham in 2019.