

# **Some criminal (if not yet criminalised) misapplications of “science”, logic, and power illustrated from the field of early childhood education.**

**John Raven**  
30 Great King St.,  
Edinburgh EH3 6QH

Website: [www.eyeesociety.co.uk](http://www.eyeesociety.co.uk)

**Version date: 12 March 2018**

**Warning: Contains material which some will find disturbing.**

In this article I argue that the majority of studies purporting to offer evaluations of educational policies and programmes are seriously misleading; indeed, viewed from a wider perspective, they cannot be considered to constitute good science. As a result, they lead to, or support, policies which have many harmful consequences. These studies, and the policies associated with them, must therefore be considered unethical<sup>1</sup>. Worse, the failure of the researchers concerned to draw attention to the limitations of their work, or challenge the policies based upon them, must be considered unprofessional and unethical. This does not mean that those studies tell us nothing ... simply that they are not fit for purpose. Conversely, the thoughtways and social practices they recursively cement need to be radically reconsidered.

One further implication of the defects in the studies reviewed is that, as with the “replication crisis”, there are very few conclusions that can be accepted uncritically. This has pervasive implications for many of those offering conventional courses in psychology.

## **Introduction and Overview**

Whilst writing what became an essay entitled *Problems with “Closing the Gap” Philosophy and Research*<sup>2</sup> \* I became increasingly shocked at the profusion of misapplications of “science”, the willingness of scientists to submit to governments research reports which were seriously misleading and had many damaging consequences, the pervasive evidence of researches’ and policy makers’ inability to reason logically, and authoritarian implementation

---

\* In overstated form, that essay reveals, among other things, that approximately one third of pupils are seriously damaged by schools, that the system as a whole fails to attend to its main goals (which have mainly to do with helping pupils to develop and gain recognition for their particular talents), that the main evaluations of school effectiveness are grossly misleading and, as a result, unethical, that there is a massive waste of resources on intervention programmes in homes and schools that essentially don’t work, and that there is alarming state intervention in people’s homes on the, as it turns out, unsubstantiated grounds that it will help to promote their children’s cognitive development and educational success; further these interventions raise serious neglected concerns having to do with human rights and respect for the values and priorities of the parents and pupils concerned.

of policies which their instigators insisted on believing to be good and right despite the damage they did to many of the intended “beneficiaries”. Beyond this, there was the pervasive destruction of professionalism via such things as the preparation of huge (eg 650–page) manuals prescribing what those concerned<sup>3</sup> must – against the threat of such punishments as being debarred from their professions – do under in endless bureaucratically-defined circumstances, regardless of the particular context and the consequences for their clients. And then there was the almost unbelievable failure of researchers and professionals to challenge extreme authoritarian interference in people’s lives (“for their own good”) despite the fact that these legally enforceable commands often involved the negation of human rights.

It is the purpose of this article to provide illustrations of these things, explore how they came about, and argue that they amount to nothing less than criminal misapplications of science, logic, and authority.

### **Scientific failures.**

The chronicle of scientific failures is almost unending.

To take a few examples, endless researchers have published evaluations of educational policies and procedures which fail to investigate or report numerous negative effects of those policies and programmes. For example, there are tens of thousands of studies of “school effectiveness” which fail to provide comprehensive evaluations of those policies and programmes or even report their deficits. Many of them focus exclusively on performance on traditional (actually invalid) tests in such areas as reading or science while failing to investigate or report on progress to, or away from, what are considered to be the main goals of education – ie helping children to develop and get recognition for their own particular talents and nurture such qualities as initiative and problem solving ability. Worse, they fail to assess the *destructive* effects of such programmes – the development of trained incapacity, gross feelings of failure and lack of worth, suicidal tendencies among those who are exposed to them and resentment and alienation among those who are debarred. Many readers will argue that this is entirely acceptable; that is the way science advances: After all, one can – and must – deal with only one thing at a time. But consider this: “Science” is thought to be concerned with objectivity. But how can a study which both reports only a few (inadequately measured) outcomes of a process and fails to report a whole range of damaging outcomes be considered objective, let alone scientific? How can it be considered ethical?

It turns out that this oversight arises mainly, although by no means entirely, from researchers’ willingness to accept the Queen’s shilling and work only on projects funded via government agencies’ “calls for proposals” to undertake specific studies and accept contractual arrangements which explicitly (i) deny them the right to study or report on issues not specified in those contracts and (ii) require them to obtain government approval for everything they propose to say<sup>4</sup>. This may be considered merely a failure to exercise professional integrity and acceptance of the need to earn a living somehow, anyhow. But hear this: I have not come across a single case in which the researchers’ concerned have underlined this gross scientific failure in the discussions of the limitations of their work that most professional bodies require them to include in their reports. Given that, as mentioned above, the result is that they do enormous damage to many, if not all, pupils and society, these studies must not only be considered unscientific, unprofessional, and unethical but also seen to constitute what are, in effect, criminal misapplications of science.

Other serious scientific failures are to be found among the almost endless researches supposedly reporting on the effects of home environment on educational performance and other psychological characteristics. Two aspects of these studies boggle the mind. First, there are virtually no comprehensive studies of home environments – of the ways parents and children interact with, and mutually affect, each other and the environments in which they live. Given that it would require a radical change in approved methodology to conduct such studies, this is perhaps not surprising. What is more surprising is that there are virtually no studies of the variance in parents’ and children’s values and priorities and their differential effects. Indeed “working class” values are not merely almost entirely ignored but also, somehow, treated with contempt. These unmentioned values, priorities, and thoughtways are somehow seen to be in need of “correction” via parental re-education programmes and stamped out.

The second amazing thing is that the outcomes studied consist almost entirely of things like “cognitive ability” (measured in ways which do not justify the application of this high-sounding and widely-encompassing term) and school performance defined in terms of test scores. In place of comprehensive studies of home environments, how they operate, and their differential effects on multiple outcomes in different school environments we find tick-box assessments of aspects of home environments deemed important from the point of view of narrowly conceptualised “cognitive development” – with “cognitive development” itself often defined as performance on “academic” tests. There are virtually no studies of effects on such things as self-confidence, problem-solving ability, or the ability to put people at ease ... let alone such things as toughness and strength (which happen to be greatly valued by some parents and children). The result is a collection of entirely one-sided and grossly misleading researches.

At this point, many readers will be thinking “What’s wrong with this? This is the way science works. It is entirely appropriate, indeed necessary, to study one thing at a time with whatever tools and methods that are available and leave someone else to venture into the other areas”.

In reality there are many things wrong with this viewpoint. In the first place, it can be used to legitimise some absolutely unjustifiable policies. More seriously, it stems from a failure to question the assumptions of reductionist science (which will be discussed more fully below) and the serious scientific and interpretational errors which such neglect brings with them. It reveals a failure to reflect on the nature of “science” itself. And it legitimises failure to consider the ethical implications of what is being carried out in the name of science and thereafter translated into practice via massive mandatory government programmes.

And then there is the conceptualisation and measurement of the educational/schooling process itself. Although a number of researchers have established that it is the wider ethos of the school which has most influence on the most important outcomes of the educational/developmental process, this observation is entirely lost in the mountain of studies which purport to offer insights into how to improve “school effectiveness”. The number of studies which have made a serious attempt to conceptualise and index what, in *Problems with “Closing the Gap” Philosophy and Research* and elsewhere, we termed *developmental environments* can be counted on the fingers of one hand. Instead we get treated to analyses seeking to elucidate the differential effects of such things as setting vs streaming, length of school day, time spent on homework, teacher qualifications, and tick-box assessments of such things as the number of questions teachers ask of the pupils in their classrooms. At what level

can studies based on such crude measures of what are essentially administrative arrangements really be claiming to offer scientific insights into the way forward in education?<sup>5</sup>

Then there is the quality and range of the outcome measures themselves. At one level many of the tests used have been developed following currently approved and promoted procedures to assess internal consistency, sometimes even involving Item Response Theory<sup>6</sup>. But, there is a pervasive tendency to fail to ask whether these tests as valid measures of the broader constructs ... such as “reading ability” or “scientific ability” ... implied by such terms. Do the tests designated as measuring “reading ability” really assess the ability to scan printed material for evidence relating to one’s purposes or explore through “lateral thinking” the implications of statements that are made ... sometimes only in passing? Do the tests of “scientific ability” index such things as the ability to problematize, seek evidence, and draw conclusions about everyday matters? In reality most of these tests measure knowledge of arbitrary snippets of information which are unrelated to the pupil’s lives or careers. And even such knowledge has a half-life of a year (pupils forget 50% after one year and 50% of the remainder after another year ... and so on). And so one has a huge superstructure concerned with teaching and assessing knowledge of little actual value using tests of which have little predictive validity outside the school system<sup>7</sup>. All of this would be of purely academic interest were it not for the fact that these tests, embedded in a wider social system, largely control what teachers teach and pupils strive to learn.

In any case, comprehensive evaluation requires assessment of *all* outcomes of an intervention and what happens in its absence. Yet there are no good measures of most of the likely outcomes – self-confidence, trained incapacity, awareness of particular strengths – and nor could there be given the acceptance of, and failure to challenge, mainstream thinking in psychometrics and measurement theory as a whole<sup>8</sup>. More generally, these observations point to the need to tailor the measures that are used to an understanding of the process that is to be evaluated and its likely effects (both positive and negative). But developing an understanding of the processes and their likely effects is not generally seen as a key component in the development of an appropriate methodology for use in the study. Perhaps one reason for this is that adopting such an approach quickly leads one to the conclusion that one will need to develop new measures – actually indices rather than measures in the traditional sense – to reflect the outcomes of the process. It cannot be done with off-the-shelf tests. Yet no time has been budgeted for this process<sup>9</sup> and, in the end, its adoption only leads to further difficulties because the resulting indices can then be discredited as not having been properly validated ... even though the processes generally used to acquire such authenticity are themselves often more than questionable<sup>10</sup>.

But there is also a more specific methodological problem. The studies conducted by the IEA<sup>11</sup> (a precursor to PISA) were promoted on the grounds that they would make it possible to understand the relative merits of the different educational processes and procedures employed in different countries. But the development of the tests was handed over to experts with international reputations ... especially expertise in Item Response Theory. So far so good (in a kind of a way). But there was a catch. To “scale” in all the 26 countries involved it was necessary to eliminate items which “worked” in only one or two contexts. That is to say, it was necessary to eliminate all the items which might have made it possible to document the effects of idiosyncratic educational practices used on only one or two countries<sup>12</sup>. So the studies *could not* do what their promoters claimed they would do. Ironically, they could do little more<sup>13</sup> than advance what their promoters specifically claimed they were *not* designed to do ... viz to promote an International Olympics in education. This, along with similar

Olympics within and between schools within countries, has subsequently been used to help impose “neo-liberal” policies (better characterised as Social Darwinistic policies designed to eliminate the weak) on pupils, teachers, and schools alike.

As if all these were not sufficient grounds for concern about the misapplication of science there remain two more issues.

One has to do with the misapplication of group tests in the assessment of individual growth and development and the evaluation of interventions.

In the course of multiple-talent and competency-oriented, project-based, educational programmes designed to promote the development of multiple and alternative competencies<sup>14</sup> and programmes of independent studies<sup>15</sup> pupils and students are expected to, and do, develop in different directions. The same applies to health care in general although it is easiest to illustrate the problem from psychotherapy where different clients develop and decline in multiple alternative ways in different directions<sup>16</sup>. The same applies in home and family-based education. Clearly the application of an arbitrary selection of off-the-shelf measures cannot do justice to the situation. Indeed, Lester<sup>17</sup> has categorised the notion of assessing self-managed learning (viz the most important learning of all) as an oxymoron. But whatever about that, the application of the most widely accepted procedures cannot result in anything but meaningless and misleading results. In other words it is grossly unscientific and the effects it has on practice unethical.

Despite these observations, one can still envisage situations in which one might be interested in measuring the differential effects on different people of different aspects of an intervention on a particular outcome. In fact there are tens of thousands of studies which purport to do this. On the face of it, nothing could be simpler – for example: Test a sample of eg high and low ability students before and after an intervention and compare the relative gains. Quite apart from the comprehensiveness issue, this procedure is actually fraught with difficulties. Most tests (eg those based on “Likert” methodology) make no claim to yield interval scales. Consequently, the same difference between two scores at different points in the scale does not imply the same difference in underlying ability... especially when comparing those with high and low scores. And, even with tests which conform to Item Response Theory, a difference score means different things depending on the absolute difficulty of the test, the shape of the test characteristic curve, and the ability range in which the difference is measured<sup>18</sup>. Thus, even by simply changing the difficulty of the test used, the relative gains of high or low scorers (or those studying in different kinds of programme) can be dramatically changed, even reversed. This calls the scientific status of many of the thousands of studies which dominate the literature into serious question.

Given that we have now seen that the field is permeated by unscientific, unjustifiable, and misleading studies, many of which have destructive consequences, one must ask how all this comes about: How have well-intentioned scientists been so widely embroiled in the process without a explosion of concern?

I will return to this in a moment. But here is a related question which has been highlighted by this review: How does it come about that concern with human dignity and well-being has been so widely corrupted into mandatory requirements to do certain things regardless of what the individual in question might choose to do?

How does it come about that a right to education has been corrupted into a mandatory requirement to attend school<sup>19</sup> however unsatisfactory that schooling may be? How does it come about that a right to a happy home life has been corrupted into a duty for parents to conform to a long list of bureaucratic specifications enforced by state-appointed guardians? How does it come about that a right to life has been corrupted into a mandatory requirement to stay alive however unsatisfactory that life may be? How does it come about that a right to “care” has been corrupted into a duty to submit to “care regimes” which are often anything but caring?

I will discuss these important research questions more fully in a later section. But first let us return to the question of how it has come about that the work of so many scientists has turned out to be so unsatisfactory once viewed from a wider perspective?

*The generation of misleading research through the funding and publication process.*

Although it is not so much a scientific failure *per se*, there has been, among researchers, a widespread willingness to apply for, and accept, research funding which precludes the discovery and publication of information which might challenge the funder’s (scientific and political) prejudices. The best known examples of this are associated with the work of Nutt<sup>20</sup> in relation to the effects of recreational drugs and criminalisation of their possession<sup>21</sup>.

The process operates as follows: Most researchers are now<sup>22</sup> obliged to seek funding through a process whereby they respond to government “Calls for Proposals”. These specify the issues to be investigated, the methods to be used to investigate them, and the way the questions to be investigated are framed. This enables government agencies to select the research to be funded and enables them to avoid funding research to investigate things single-minded politicians do not want to know. And the terms of the contracts frequently explicitly forbid the researchers from pursuing issues other than those laid down in the Call for Proposals. Beyond that, these contracts often give the funding agency the right to actually alter the figures (numbers) to be reported as outcomes of the research (and have in practice been known to do just this). A few glaring examples have surfaced in the press, but the process is much more widespread than might be assumed ... with researchers justifying their behaviour by saying that it is not only their own but also their collaborators careers which would be jeopardised if they protested. Many contracts also require the researchers to get government approval for anything they wish to publish. The process thus results in research which may be said to have been “designed” to get results which support government perceptions and policies rather than the kinds of open-ended research which might offer a basis for alternative policies. Although the results of such research are presented as contributing to evidence-based policy, they are thus best characterised as contributing to policy-based evidence.

The effects are further exacerbated via publication processes which require researchers to submit their proposed publications for peer review. Such peers are often reluctant to agree to publications which challenge conclusions they have drawn from their own research ... and upon which their careers depend. Although, as has been mentioned, the most notorious examples of this process are to be found in the U-tube videos posted by Nutt, it is also evident from the way in which studies and discussion of the wider goals of education have virtually disappeared from the educational research literature since the early 1980s ... which coincides with the introduction of the national curriculum and continuous testing via

standardised tests. School effectiveness (and its assessment) has come to mean success in these terms and nothing else.

The way in which these two processes combine to exert a stranglehold on research is further amplified by the increasing government-mandated pressure to “publish or perish”. This offers yet another example of the brutal imposition of social Darwinism in the guise of (neo) liberalism. In the name of promoting efficiency and accountability in Higher Education the Government has imposed a series of attempts to improve “Quality” through what were first called Research Assessment Exercises and later a Research Excellence Framework. These not only require staff to raise funds (largely through the processes described above) but also to, for example, publish 7 papers a year in “high impact” journals ... ie those which are widely cited ... ie those which serve the “generate papers which will least challenge the peer-evaluation process and get cited by those peers” process – thus forming a self-reinforcing cycle.

*(Unthinking) acceptance of Reductionist Science.*

Contrary to the most widely accepted assumptions about what needs to be done to advance understanding via experimental, theory-oriented, science (which revolve around studying the effects of varying something [a single variable] on a pre-specified outcome [assumed to be important from a theoretical or practical point of view]), my claim is that **to be acceptable as a scientific evaluation of an intervention it is necessary for it to strive to be comprehensive.**

By this I mean that it is necessary for the researchers concerned to strive to document *all* the personal and social (what is good for the individual may be bad for society), short and long term (what is good in the short term may be bad in the long term), intended and unintended, desired and desirable, and undesired and undesirable effects of an intervention (which may itself, as discussed below, amount to much more than whatever was intended, let alone, reported by its instigators) on different people in different social contexts. Otherwise the reported results corrupt the advance of understanding.

In other words, *the quality of an evaluation* is to be judged more in terms of its **comprehensiveness** – ie the extent to which it offers a rough fix on **all** important inputs and outcomes – than in terms of the accuracy of its assessments relating to any one of them.

Focussing only on the magnitude of *intended* effects (as legitimised by conventional, over-simplistic, images of theoretically-based “experimental” science) may lead to failure to study highly undesirable unintended consequences.

This can not only lead to disastrous policies; it also confuses the issues ... i.e. undermines understanding ... and thus constitutes bad science.

Well-known examples of studies legitimised via acceptance of reductionist science include failure to study the effects of fertilisers and pesticides on such things as the future fertility of the soil (itself an emergent property stemming from the complex interactions between multiple complex organisms), long term effects via the food chain on a wide range of species (including ourselves), and the effects on the diversity of species living in complex symbiotic relationships within human beings<sup>23</sup>.

Actually, to advance understanding, it is not only the *outcomes* which need to be comprehensively documented. The actual (as distinct from the stated) nature of the “intervention” needs to be scrutinised. For example, experimental variations in health care in different contexts (eg homes vs hospitals that themselves vary greatly in the wider policies they pursue) may bring with them overlooked variation in what are taken to be the inputs. For example, patients involved in experimental trials may get more attention than those to whom the drug will, in due course, be administered routinely. In education, participation in a study nominally designed to enhance “cognitive ability” may bring with it all sorts of benefits not available to a control group. And the varying context of that intervention may result in very different outcomes. Thus a programme to nurture “cognitive development” will have very different effects in homes which value such outcomes and in those in which parents disparage it.

Nor is this the end of the matter. The incorporation of single-factor thinking into policy-making inhibits any tendency to set up a variety of experiments to cater for people who have different priorities to one’s own and to assess the effects of alternative policies<sup>24</sup>.

In reality both of these observations illustrate a wider issue, namely the, actually horrendous, pervasive and pernicious effects of neglecting systems thinking in policy making and science that are discussed in Raven (2016) but omitted here for space reasons.

Despite the need to limit our discussion, it may nevertheless be useful to give one example of the effects of neglecting systems thinking in “management”.

To many people it seems obvious that the performance of systems can be improved by setting “targets”. Yet, counterintuitively, in reality, the setting of “targets” always makes things worse<sup>25</sup>. This is because they deflect people’s attention away from the goals the system was intended to achieve and the multiple things which would need to be done to achieve them and instead lead them to invent ways of meeting the mandated targets without doing what would actually need to be done to achieve the system goals effectively<sup>26</sup>. The effect of setting very narrow targets (eg performance on tests, state mandated or otherwise) and associating them with serious rewards and punishments for pupils, teachers, and administrators alike is there for all to see in “education”.

All of this may seem obvious. But it has not *been* obvious in the past. On the contrary, both “scientists” and “decision takers” – politicians – tend to focus on single issues, thereby overlooking the fact that, as we have seen, *single-factor intervention in poorly understood systems almost always has counterintuitive, and usually counterproductive, effects*.

Although I have already used it to illustrate general scientific failure, I will now say a little more about the effects that a preoccupation with single outcomes has had on studies of the effects of home environment on children’s development.

We have seen that psychologists have, on the whole, been preoccupied with “cognitive development” (itself poorly conceptualised and measured) and failed to document the wider outcomes (both positive and negative) of educational and developmental processes in homes and schools. At one level, it seems that the field has been captured by a single factor concept of “ability”. (*Problems with “Closing the Gap” Philosophy and Research*<sup>27</sup> is specifically devoted to generating a – mainly *sociological* – explanation of how this comes about.) But even those who do mention the wider abilities which people might possess<sup>28</sup> are hamstrung



by the fact that, as I have shown elsewhere<sup>29</sup>, the identification of these qualities requires psychologists to move away from thinking in terms of *variables* (as in physics) to a *descriptive* paradigm as in biology. This makes it impossible to identify the benefits and disbenefits of the various kinds of environment found in homes, schools, and workplaces. As a result, the thousands of studies which purport to offer insights in this area not only cannot, in any meaningful sense, claim to be scientific, but must also be considered unethical since they condemn so many people to destructive and degrading “educational” programmes and thereafter to occupations which fail to develop and utilise their abilities.

The paucity of meaningful frameworks for thinking about and describing the variety of developmental or destructive environments that exist in homes, schools, and workplaces has already been mentioned. Here it is important to emphasise that, in the light of the mountain of claims about the importance of home environments and the use of such claims to justify massive intrusive mandatory interventions in homes and schools this is not only truly remarkable but highly unethical.

In reality, as has been mentioned, not only are there virtually no studies of what actually goes on in homes and the complex, reciprocal, mutually-influencing interactions which occur between parents, children, and their environments, it would be virtually impossible to conduct them without a radical change in approved methodology. It would be necessary to develop a framework akin to that employed by some biologists to study the complex symbiotic interactions which occur between species in different habitats.

Actually, the situation is even worse than might be suspected. It emerges that most assertions about what goes on in homes are not only preoccupied with factors which are assumed to be related to “cognitive ability” (and ignore other aspects of competence), they are based on conjectural backward projections from observations made by middle class teachers and researches of the differences between the behaviour children in nursery schools!

And Rich Harris<sup>30</sup> has shown that even, at the level at which they exist, the tens of thousands of studies which purport to show relationships between (limited aspects) of home background and (grossly inadequate) measures of “ability” are seriously flawed because they fail to investigate the way the observed relationship comes about. When this is done, it emerges that it is mainly the variance in children’s behaviour which causes the variance in parents’ behaviour rather than the reverse. In short this vast array of studies mainly provide examples of the classic methodological error of interpreting a correlation as demonstrating a causal relationship and then misinterpreting the direction of that supposed causal relationship<sup>31</sup>. What then happens is that error spreads like wildfire through the writings and conferences of thousands of others who are not given to critical thinking and are, in any case, deterred from exposing the error by the publish-or-perish-via-positive-peer-review system and sound-bite based communications.

As if all this were not enough, researcher after researcher has been satisfied to interpret the correlation between children’s test scores early in life with those same children’s scores five or more years later to mean that the first causes the second and thus that intervention early in life – especially with the “less able” or those from certain backgrounds – will collapse the correlation. And, even if it does not do that, it will increase everyone’s scores so that they will all get jobs and do well in life. Quite apart from the fact that, as I show in *Problems with “Closing the Gap” Philosophy and Research*, there is not only no evidence to support these assertions but plenty of evidence to the contrary what we should be most concerned about

here is the widespread failure to question the *logic* of such assertions; the failure to demand an elucidation of the alternative mechanisms whereby the observed correlation may come about<sup>32</sup>.

### **Failure to think logically**

The error of inferring causal relationships from correlations<sup>33</sup> recurs repeatedly in the literature reviewed while writing *Problems with “Closing the Gap” Philosophy and Research*.

Among the most striking is that of concluding from the observed correlation between educational attainment and whether or not people get jobs is that, if everyone gets more education, everyone will get jobs. This arises from neglecting the norm-referenced nature of both the dependent and independent variables. The relationship persists even if everyone gets more education. All that happens is that everyone stays in the system longer (admittedly thereby lowering national levels of unemployment by keeping people out of the job market and creating jobs in the educational system itself) and employers raise entry requirements without there being any change in the competencies actually required in the jobs.

A similar error occurs in the course of demonstrating that “remedial” programmes targeted at “those with special needs” (marginally) improve their scores while failing to note that those not included in the programme thereby necessarily come to occupy lower places in the pecking order. That is the way norm-referenced systems work. Apart from any genuine gains in competence that may have been gained (which are hard to measure and, as a number of researchers<sup>34</sup> have shown, few in number) the systems benefits are zero.

It is claimed that regular mandatory testing will motivate lower ability groups to work harder. In reality, because of the norm-referenced nature of the system, this simply reinforces the “less able’s” perception that they are stuck: life is hopeless. There is no hope of getting out of the low status positions in which they find themselves no matter how hard they try. They learn that they had better accept a life of being assigned to demeaning positions and being pushed about by hoards of remediators, bureaucrats, and do-gooders.

Then there is the case of prescriptive remediation. I do not refer here to the errors which arise from neglecting the norm-referenced nature of the system but from the belief that, for the sake of “efficiency”, prospective providers should tender to provide pre-specified services so that managers can select the cheapest.

The error is the belief that pre-specified services ever could cater for the wide variety of problems that are assigned to certain categories (such as “dyslexia”) on the basis of superficial similarities.

Following through with the “dyslexia” example, prospective providers are asked to tender for providing services that will fix the rag bag set of problems that are so categorised<sup>35</sup>. For this process to work, it has to be assumed that all will require approximately “the same” treatment – otherwise it would not make sense to compare tenders. Nothing could be further than the truth – and the most effective “treatments” would require reform of the “educational” system itself. The problem is even clearer in the Health service where commissioners require prospective providers to tender for providing pre-specified (and unevaluated) services at certain points as patients move along pre-specified “paths”. As Seddon<sup>36</sup> has shown,

precisely because they are not tailored to the patient's particular needs, these rarely work. The result is that patients re-present with the same symptoms and complaints as they had at the beginning of the process. This greatly inflates the apparent demand for the ineffective service ... Seddon calls it "Failure demand".

The explanation of this process probably lies in a deep-seated (authoritarian) belief that the way to avoid failure is to pre-specify – ie incorporate into manuals through a process best described as "manualisation" – the procedures that are to be followed in pre-specified situations. The result is to destroy professionalism<sup>37</sup>.

But that is not the end of the matter. There has been a dramatic failure to study the *systems* causes of the problems. Thus "dyslexia" chiefly arises from failure of the system to nurture and recognise the wide variety of talents and deficits that exist and a preoccupation with a limited notion of "reading ability" measured in pre-specified (and largely invalid) ways<sup>38</sup>. In reality, most of the problems the "health" services seek to cater for arise from the way we organise work, society, and "benefits".

What I am saying here is that these deficits arise from a failure to think comprehensively, logically, about these problems, their causes, and their remediation. More specifically, they reflect the neglect of systems thinking.

### **Criminal intervention in people's life styles and homes.**

Many mandatory, state-wide, policies which end up enforcing activities which have consequences which must, in many cases, be regarded as criminal have already been mentioned. These consequences go far beyond, but include, denial of human rights.

In order to understand this we must first note that these policies were often initiated on the basis of some observation of a matter of genuine concern, such as the consequences of illiteracy or child neglect. However, the services that are offered then cease to be services *on offer* but get translated into commands that everyone behave in certain ways or conform to certain requirements regardless of the values of those concerned or the particular circumstances in which they find themselves. These commands are often backed by threats of punitive sanctions ... including the threat of criminalisation.

It is in this way that the desirability of everyone having access to some form of education (a term which implies nothing more than participation in some programme which will "draw out", or nurture, the individual's particular talents) gets translated into compulsory<sup>39</sup> attendance (enforced by an army of heavy-handed school attendance officers) at schools which are required to follow a narrow, centrally specified, curriculum. Progress toward these mandatory goals is then assessed using norm-referenced tests having little construct or predictive validity. The inevitable failure designed into norm-referenced testing (by definition, 50% fail in the sense of being below average) is followed by a range of severe mandatory punishments for pupils, parents, teachers, and schools alike. The resulting destruction of many pupils lives cannot be considered anything short of criminal ... and it is hard to see how the anguish caused among many teachers and parents ... never mind the social destruction wreaked by a divided society ... can escape a similar blandishment.

More generally, the over-riding of many pupils' and parents' values and preferred lifestyles amounts to a severe infringement of their human rights.

To turn to another example. Denial of opportunity to live a high-quality life (defined in one's own terms) seems an affront to civilised society. But what then happens is that bureaucrats set about defining what constitutes a high quality life and then command its imposition on everyone via, in the case of the Scottish *Children and Young Peoples Bill*, armies of state-appointed guardians wielding 36 pages of tightly printed tick box questions<sup>40</sup> relating to the quality of children's home environments, the criminal records and financial status and lifestyles of the parents, and the children's own values and behaviour. These assessments of the children include such things as whether or not they have a positive attitude to school and "appropriate attitudes" toward their own sexuality.

The questionnaires are overwhelmingly informed by middle class values – middle class values and thoughtways that are then embedded in the research of young middle class researchers with no children of their own.

And the conclusions drawn from these assessments by these mandatory state-appointed guardians are then reviewed and enforced by mandatory committees the members of which (unlike the parents accused of failing to promote wellbeing of the prescribed type in the prescribed ways) have access to all the family's financial, health, educational, and criminal records and have powers to have children taken into care<sup>41</sup>, compel the parents (despite its costs and intrusion into their lives) to attend compulsory re-education programmes, or, in the last resort be subject to fines and imprisonment. And all without any provision for representation on behalf of the parents or children concerned.

Clearly these mandatory arrangements embody a gross disrespect for personal values and human rights.

They appear to be based on a pervasive human predisposition to believe that one has right to impose, by force if necessary, whatever one believes to be good and right on others regardless of the consequences for those concerned. This predisposition seems to extend to a frightening willingness to condemn and ostracise people with other values who make themselves visible through the social media, and, not necessarily more seriously, a historical willingness to assign others to concentration camps, torture, or burning at stake for "inappropriate" political or religious beliefs.

It also extends to pervasive acceptance of the notion that it is appropriate in a democracy to take decisions which are binding on all despite the variation in opinion and the inappropriateness of the proposed actions. Indeed the essence of "democracy" is largely seen as inhering in the voting process rather than in a process which enable people who have different priorities to lead their lives in their own way.

More generally, one may comment that the imposition of the neo-liberal notion that there are some able children who must be sorted out and promoted for the good of society as a whole ... policies which have wreaked such destruction in schools and society ... seems usually to have been introduced without any awareness of possible alternatives ... that is to say on the basis of the hegemony which this way of thinking has secured. While it is true that so-called neo-liberal policies are usually, perhaps invariably, backed by force ... by the army in the case of economic policies<sup>42</sup> and threats of having children taken away, consignment to remedial re-education programmes, and imprisonment in the case of schooling ... the

hegemony of neo-liberal thoughtways play a much more important role. They inform policy in almost every domain.

In passing, we may note also the tendency of those who wish to promote authoritarian policies to seize on research – however flawed<sup>43</sup> – to justify their proposals.

### **Criminal misuse of authority/power.**

I have touched on examples of what can only be described as the criminal misuse of authority throughout this article, but will now bring a number together in a bulleted list.

As I see it, what is in effect the criminal misuse of authority occurs in the:

- Constriction of university research funding (almost a prerequisite to advancement in academe) mainly to that available via government “calls for proposals” to undertake tightly prescribed and monitored research, incorporating specific proscription of enquiry into issues *not* specified in the research (thereby eliminating the traditional role of the university – which was to engage in free-ranging enquiry<sup>44 45</sup>).
- Restriction of publications from such research to those approved by government.
- Inclusion of a right to actually alter figures in the reported results of such research.
- The elimination of academics’ time to think via the pressures generated by the requirements of Research Assessment Exercises (Research Excellence Frameworks).
- Constriction of research to small scale experimental and pilot studies instead of tackling of wider issues (see endnote above).
- Elimination of challenge to narrow and conventional perspectives by the peer review process required for publication in “high impact” journals as required by the Research Excellence Framework.
- Enforcement of commands to attend schools (even though that process may be highly destructive) via an army of enforcement officers affecting pupils directly and their parents.
- Introduction of mandatory curricula concentrating on imparting and testing narrow snippets of irrelevant knowledge and neglecting the wider competence goals educators could potentially pursue ... and imposing this framework even on Home Educators.
- Imposition of mandatory national testing programmes at regular intervals (to reinforce pupils’ knowledge of their status in the pecking order) using tests constructed according to principles of dubious merit.
- Using the results of these tests to orchestrate Educational Olympics within classrooms, between schools, and between countries ... Olympics which have few winners but millions of failures i.e. Social Darwinism.
- Introduction of armies of inspectors with extraordinary powers to intervene in homes and schools and punish (even via punitive “remediation” often involving giving up otherwise free time) pupils, parents, teachers and head teachers alike.
- Mandatory bureaucratic generation of rules and manuals of procedure to control and prescribe the behaviour of children, parents, teachers, and social workers ... all coupled with the generation of endless training programmes to teach all concerned the rules embedded in the Manuals. In short, mandatory destruction of professionalism.
- Requirements to seek tenders for providing centrally-specified, assumed to be routine, services – a process known as “commissioning” in the Health Service – when what is

required is a range of client- and situation- specific services (hence generation of “failure-generated demand” for more of these (ineffective) services.

- Stetting targets (test scores; time to achieve specified outcomes; number of pupils enrolled etc.) the achievement of which deflect the attention of those concerned from the goals of the system.
- Mandatory intervention in homes to impose values which may be foreign to the parents and families concerned.
- Mandatory infringements of human rights via data sharing (health, crime, income, home environment assessments) among armies of inquisitors.
- Removal of children and parents from homes and replacement by regimes of care (personal or institutional) which often turn out to be anything but careing.
- Corruption of *rights* (eg to education, well-being) into *requirements* (eg to attend schools (however bad); to provide specific types of home environment; to display “appropriate attitudes toward own sexuality”; etc.) accompanied by heavy-handed monitoring followed by punitive sanctions.

At this point, it is perhaps appropriate to, once again, underline the pervasive influence of neo-liberal thoughtways – ie that what is important is competitive success at centrally decreed tasks whatever those tasks may be and regardless of who it is that has the right to specify those tasks. The result is nothing less than the brutal imposition of Social Darwinism.

The source of this pervasive fascism – the belief that one has the right to impose on others, by force if necessary, thoughts and behaviours that one believes to be good and right regardless of the consequences for those individuals and society, and the implied denial of the right of individuals to take decisions for themselves – merits investigation in its own right<sup>46</sup>.

## REFERENCES

- Bailey, D.H., Duncan, G.J., Watts, T., Clements, D.H., & Sarama, J. (2018). Risky Business: Correlation and Causation in Longitudinal Studies of Skill Development. *American Psychologist* 73, 81-94.  
<http://dx.doi.org/10.1037/amp0000146>
- Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, 2(1), 67-90.
- Deming, W. E. (1993). *The New Economics for Industry, Government, and Education*. Cambridge, MA: Massachusetts Institute of Technology.
- Elliott, J. G., & Grigorenko, E. L. (2014). *The Dyslexia Debate*. Cambridge, UK: Cambridge University Press.
- Fink, B. (2016). How Neoliberalism Got Organized: A Usable History for Resisters, With Special Reference to Education. *The Good Society*, 25, 158-171.
- Flynn, J. R. (2000). *How to Defend Humane Ideals*. Nebraska: University of Nebraska Press.
- Hamilton, D., Jenkins, D., King, C., MacDonald, B., & Parlett, M. (Eds.). (1977). *Beyond the Numbers Game*. London: MacMillan Education.
- Happer, R. (2017). This is about their truth. *The Psychologist*, December 2017, 52-54
- Harris, J. R. (2006). *No Two Alike: Human Nature and Human Individuality*. New York: W. W. Norton.
- Hattie, J.A.C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge, Taylor & Francis.
- Hope, K. (1984). *As Others See Us: Schooling and Social Mobility in Scotland and the United States*. New York: Cambridge University Press.
- Hunter, J. E., & Hunter R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Kazdin, A. E. (2006). Arbitrary metrics: Implications for identifying evidence-based treatments. *American Psychologist*, 61, 42-49.
- Klein, N. (2007). *The Shock Doctrine: The Rise of Disaster Capitalism*. London: Penguin Book

- Lester, S. (2001). Assessing the self-managing learner: A contradiction in terms? Chapter 26 in J. Raven & J. Stephenson (Eds.), *Competence in the Learning Society*. New York: Peter Lang.  
<http://eyeonsociety.co.uk/resources/CILS-chapter-26.pdf>
- Maxwell, J. N. (1969). *Sixteen Years On*. Edinburgh: Scottish Council for Research in Education.
- O'Reilly, D. (2001). Competence and incompetence in an institutional context. Chapter 22 in J. Raven & J. Stephenson (Eds.), *Competence in the Learning Society*. New York: Peter Lang.
- Prieler, J. A., & Raven, J. (2008). The measurement of change in groups and individuals, with particular reference to the value of gain scores: A new IRT-Based methodology for the assessment of treatment effects and utilizing gain scores. In J. Raven & J. Raven, (Eds.), *Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*. Unionville, New York: Royal Fireworks Press; Edinburgh, Scotland: Competency Motivation Project; Budapest, Hungary: EDGE 2000. Also available at <http://eyeonsociety.co.uk/resources/UAIChapter7.pdf>
- Raven, J. (1984). *Competence in Modern Society: Its Identification, Development and Release*. Unionville, New York: Royal Fireworks Press. (First published in 1984 in London, England, by H. K. Lewis.)
- Raven, J., Johnstone, J., & Varley, T. (1985). *Opening the Primary Classroom*. Edinburgh: Scottish Council for Research in Education.
- Raven, J. (1991). *The Tragic Illusion: Educational Testing*. New York: Trillium Press. [www.rfwp.com](http://www.rfwp.com) (now available from the author at 30, Great King Street, Edinburgh EH3 6QH.)
- Raven, J. (2008). The need for, and development of, the SPM Plus. In J. Raven & J. Raven (Eds.) *Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*. Unionville, New York: Royal Fireworks Press; Edinburgh, Scotland: Competency Motivation Project; Budapest, Hungary: EDGE 2000; Cluj Napoca, Romania: Romanian Psychological Testing Services SRL. (Chapter 3, pp. 103-112).
- Raven, J. (2014). Dyslexia - getting it wrong. *The Psychologist*, 27(11), 809.  
<http://eyeonsociety.co.uk/resources/Dyslexia.pdf>
- Raven, J. (2017) Problems with Closing the Gap Philosophy and Research, *Baltic Humanitarian Journal*, 3, 252-275. Full version at: <http://eyeonsociety.co.uk/resources/Closing-the-Gap-2017-V2.pdf>
- Rothschild, Lord. (1982). *An Enquiry into the Social Science Research Council*. London: HMSO.
- Schmidt, F. L., Oh, I.S. & Shaffer, J. A. (2016). The Validity and Utility of Selection Methods in Personnel Psychology: Practical and Theoretical Implications of 100 Years of Research Findings. Unpublished, available from the authors.
- Seddon, J. (2008). *Systems Thinking in the Public Sector: The Failure of the Reform Regime ... and a Manifesto for a Better Way*. Axminster, UK: Triarchy Press.
- Schön, D. (2001). The crisis of professional knowledge and the pursuit of an epistemology of practice (Chapter 13). In J. Raven & J. Stephenson (Eds.), *Competence in the Learning Society*. New York: Peter Lang. Also available at <http://eyeonsociety.co.uk/resources/CILS-chapter-13.pdf>
- Schweinhart, L. J., & Weikart, D. P. (1997). The High/Scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly*, 12, 117-143.
- Shiva, V. (1998). *Biopiracy: The Plunder of Nature and Knowledge*. London: Green Books.
- Scottish Government. (2014). *Children and Young People (Scotland) Act, 2014*. Norwich, England: TSO (The Stationery Office). [http://www.legislation.gov.uk/asp/2014/8/pdfs/asp\\_20140008\\_en.pdf](http://www.legislation.gov.uk/asp/2014/8/pdfs/asp_20140008_en.pdf)
- Spearman, C. (1927). *The Abilities of Man: Their Nature and Measurement*. London, England: MacMillan.
- Stephenson, J. (2001). Inputs and outcomes: The experience of independent study at NELP (Chapter 21). In J. Raven & J. Stephenson (Eds.), *Competence in the Learning Society*. New York: Peter Lang.  
<http://eyeonsociety.co.uk/resources/CILS-chapter-21.pdf>
- Suggate, S. P. (2012). Watering the garden before a rainstorm: The case of early reading instruction. In S. Suggate, & E. Reese, *Contemporary Debates in Childhood Education and Development* (Chapter 17). London, England: Routledge.
- Taylor, C. W. (1974). Multiple talent teaching. *Today's Education*, March/April, 71-74.
- Weerts, D.J. (2016). From Covenant to Contract: Changing Conceptions of Public Research Universities in American Society. *The Good Society*, 25, 182-208.
- Wilkinson, L., & Task Force on Statistical Inference. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54, 594-604
- Wolf, A. (1987). *Work Based Learning: Trainee Assessment by Supervisors*. Bradford, England: MSC Sales, ISCO.

## ENDNOTES

<sup>1</sup> For a remarkable discussion of such issues see Flynn (2000).

<sup>2</sup> Raven (2017), which contains links to many of the sources supporting statements made in the current article.

- 
- <sup>3</sup> Eg Teachers in their classrooms, social workers with their clients, and those commissioned to provide health “services”. These prescribe what teachers will do with *all* children in their classes, social workers shall do in certain situations regardless of the particular needs and circumstances of their clients and the context in which they are working, and how health and care providers shall classify their patients and clients and thereafter offer packaged treatments that do not, in reality, relate to their needs. All standardise provision in ways which prevent those concerned paying attention to the particular needs and circumstances of their clients and acting accordingly. In these ways they prevent those concerned exercising their skill and judgment and thus destroy professionalism. (See also Schon, 2001).
- <sup>4</sup> Although this is probably the main reason for their neglect since “neo-liberal” thoughtways overtook the academic world (see Weerts, 2016, for a discussion of alternatives), other reasons include the paucity of measures of these wider outcomes and thoughtways which inhibit the development of appropriate indices.
- <sup>5</sup> This is not to deny that some of these studies do offer important insights. For example, one outcome of the IEA and PISA studies is that, by the time children are 11 years of age, there are no differences in the much sought after areas of educational attainment between school systems which admit children from all ages from 4 to 8.
- <sup>6</sup> But even here it is not always clear that the researchers concerned have fully understood what they are doing but have, for example, simply fed their data into off-the-shelf statistical packages yielding esoteric indices.
- <sup>7</sup> Hunter & Hunter (1984), Schmidt et al (2016)
- <sup>8</sup> For a fuller discussion of this process see eg Raven (1991)
- <sup>9</sup> Stephenson (2001), Lester (2001).
- <sup>10</sup> While such observations have led many outstanding researchers (eg Hamilton et al, 1977) to, at least partially reject quantitative evaluation, and others to desert the field entirely, the alternatives are usually dismissed as “unscientific” and, as a result, unlikely to be funded by government departments.
- <sup>11</sup> International Association for the Evaluation of Educational Achievement.
- <sup>12</sup> To be fair, the IEA went to great lengths to document the different curricula that were operational in different countries and even conducted a “civics” study to document the wider outcomes of the educational processes in this area. Unfortunately, the latter ended up providing a wonderful example of the misapplication of current notions of what constitutes “science”. The initial report, looking at patterns of item statistics (as distinct from “factor scores”), documented the dramatically different patterns of beliefs which existed among the pupils from the various countries. This approach --- building up patterns from pupils’ statements treated as having some meaning --- was deemed “unscientific” because there were no “scores” ... only summarising descriptive statements. As a result that the analysis was handed over to others who submitted the data to factor analysis etc. ... and ended with a report which failed to tell anyone anything very much.
- <sup>13</sup> There is at least one important exception to this statement although its derivation, in itself, challenges much thinking about the way scientific understanding advances. Reflecting on the 800 meta-analyses he had before him, Hattie (2009) concluded that key features of the educational process cantered around such things as teacher enthusiasm and what he called “visible learning” ... teachers making the process of learning visible through their own behaviour. OK. But, from the point of view of what we are about here, what is most interesting about this is that it reverses our understanding of the processes through which science advances. These were not insights *tested in* the studies. They were insights *emerging from reflecting on* the meaning and interpretation of studies that had already been conducted.
- <sup>14</sup> See, for example, Raven et al (1985).
- <sup>15</sup> For a brief account of the work at the North East London Polytechnic see O’Reilly (2001). For the problems which the diversity of outcomes posed for evaluation see Stephenson (2001).
- <sup>16</sup> See Kazdin (2006).
- <sup>17</sup> Lester (2001).
- <sup>18</sup> Prieler & Raven (2008). (However, despite the risk of being accused of advertising, I may mention that I do know of one test to which these strictures do not apply because it has an almost *linear* Test Characteristic Curve. That test is the Raven *Standard Progressive Matrices PLUS* test. (Raven, 2008)).
- <sup>19</sup> To avoid a digression we may skip over to problems associated with home schooling at this point.
- <sup>20</sup> Nutt e.g. <https://www.youtube.com/watch?v=MRLXt1oIsqI&t=2741s>
- <sup>21</sup> For the purposes of this discussion we will ignore the diversionary nature of the “war on drugs” itself.
- <sup>22</sup> This has not always been the case and, as Weerts (2016) emphasises, changing the process is itself one of the most important priorities requiring attention.
- <sup>23</sup> See, for example, Shiva (1998).
- <sup>24</sup> Once combined with what seems to be a pervasive predisposition to believe that one has a right to impose what one believes to be good and right on others by force – and regardless of most of the consequences for people who embrace alternative values – the process leads directly to the emergence of what are, in effect, fascist policies and, in the end, Fascist regimes.
- <sup>25</sup> Deming (1993), Campbell (1979).



- 
- <sup>26</sup> See Spearman (1927) for a discussion of the effects of using conventional tests in education and Deming (1980), Kohn (2000), Seddon (2008) and Campbell (1979) for wider discussions.
- <sup>27</sup> Raven (2017).
- <sup>28</sup> Eg Spearman (1927), Taylor (1974).
- <sup>29</sup> eg Raven (1984, 2017)
- <sup>30</sup> Harris (2006)
- <sup>31</sup> This problem was discussed at some length in the report of the APA Task Force on Statistical inference (Wilkinson et al, 1999)
- <sup>32</sup> Bailey *et al* (2018) have devoted a whole article to evidence supporting this claim.
- <sup>33</sup> Coincidentally an extremely important article on this topic has appeared in *The American Psychologist*. Authored by Bailey *et al* (2018). Among other things it shows, first, that the apparent effects of early intervention are much less than is commonly thought and then that a combination of stable but unmeasured variables and generalised skill-building abilities can reproduce the apparent results of the interventions.
- <sup>34</sup> Maxwell (1969), Hope (1984), Wolf (1987)
- <sup>35</sup> Elliott and Grigorenko (2014)
- <sup>36</sup> Seddon (2008)
- <sup>37</sup> Schon (2001)
- <sup>38</sup> Raven (2014).
- <sup>39</sup> Prior to the introduction of Margaret Thatcher's GERBIL, education was explicitly the responsibility of parents who could choose whether that responsibility would be fulfilled by sending their children to school "or otherwise". Nevertheless many of those who chose the second option found were confronted by inspectors who had very limited notions of what education was about. The current situation is much more ambiguous and much more loaded toward the National Curriculum, but the option nominally still exists.
- <sup>40</sup> Known as SHANARI and GIRFEC. See Scottish Government (2014).
- <sup>41</sup> It is invariably assumed that the mandated regimes of care – whether institutional or within-family – will be better but the work of an ongoing programme of monitoring, the results of which are summarised by Happer (2017) has shown that this is far from being the case. The cases which make the headlines are *not* exceptions.
- <sup>42</sup> Klein (2007).
- <sup>43</sup> eg Schweinhart & Weikart (1997) but also Suggate (2012).
- <sup>44</sup> See Weerts (2018).
- <sup>45</sup> It is of more than passing interest to note that, in response to Mrs Thatcher's request that he write a report which would help her to close the Social Science Research Council, lord Rothschild (1982), while promoting the customer-contractor principle, both noted that social scientists were their own worst enemies because they embraced narrow academic studies and avoided wider and more socially important issues, also called for the budgeting of commissioned research to add 5% of the research *and development* costs for *scientist initiated* research. Given that the development costs of many educational projects are huge, this would amount to an enormous sum of money.
- <sup>46</sup> In this context attention may be drawn to the work of Fink (2018)