



Measuring character strengths and virtue among UK school children aged 14 and 15

David Walker

These are unpublished conference papers for the 'Can Virtue Be Measured?', held by the Jubilee Centre for Character and Values at Oriol College, Thursday 9th – Saturday 11th January 2014. These papers are works in progress and should not be cited without author's prior permission.



**School of Education
University of Birmingham
Edgbaston
Birmingham**

David Walker PhD

Measuring character strengths and virtue among UK school children aged 14 and 15

INTRODUCTION

In September 2012, the Jubilee Centre for Character and Values formed and work began on one of its flagship research projects: 'Character education in UK schools'. This project is intended to take the temperature of character education in 100 schools across the UK. The broader aim of the work is to answer the main research question: What do we know about the nature, impact and current understandings of education for character in British schools – and how can such education be improved upon in light of what we know about pupils' current states of character? So far, more than a year later, 36 schools from across the UK have taken part in this research, involving approximately 2000 pupils and 100 teachers. The data collection period will be complete by March 2014 and analysis and writing will follow until this research phase of the project makes way for more intervention focussed work in March 2015. Acknowledgements for contributions to this paper are shown in the footnote¹.

In response to the question of this conference: 'Can virtue be measured?', the present paper will set out how virtue (and character) is being 'measured' in this particular Jubilee Centre research project and report some very early and partial results. The focus of the paper is not the entire research project, but specifically those aspects of the research that relate to pupils in Year 10 (aged 14 and 15) and their teachers in secondary schools.

Recent attempts to measure virtue and character have often been designed as self-reporting measures or moral dilemma tests (See Table 1). Both of these methods have attracted some degree of criticism when they have been used in isolation: self-reporting methods for problems of social desirability and self-delusion, and moral dilemma tests have been objected to under the charge that

¹ Professor James Arthur; Professor Kristján Kristjánsson; Mr Michael Roberts; Dr Wouter Sanderse; Miss Chantel Jones; Professor Steve Thoma, Alabama University; Dr Ryan Niemiec, VIA Institute on Character.

moral reasoning does not equate to moral action. In this paper, I shall argue that a combination of different methods currently provides a promising means for the measurement of virtue (and character). In this flagship Jubilee Centre project, three methods are combined (or triangulated). These are: moral dilemmas, pupil self-reports of character strengths, and teacher interviews which include reports on pupils' character strengths. For the moral dilemmas, we have used an adapted and partial version of the Intermediary Concept Measure (ICM) for adolescents developed by Professor Thoma at the University of Alabama, USA. To achieve pupil reports on their own characters we have chosen the new shortened 96-item Values in Action Inventory for Youth measure (VIA Youth Survey) from the VIA Institute on Character in Cincinnati, USA. And, finally, in semi-structured interviews, teachers are asked to make judgements about the same groups of children surveyed by the other two measures. In this paper, I will discuss how we came to choose this approach; explain what we hope to achieve; and set out and discuss some very early and tentative results.

RESEARCH DESIGN

From the beginning, Aristotelian inspired Virtue Ethics informed our work. More specifically, the understanding that good character comprising virtue is the route to a flourishing life motivated our search in the earliest stages of our research design for an Aristotelian measure of virtue. The closest we came to this was Howard Curzer's Sphere-Specific Moral Reasoning and Theory Survey (SMARTS). Unfortunately this is not intended for the age group of present interest, and so with no off-the-shelf Aristotelian measure of virtue available to us at this time, we started to look further afield. Our philosophical grounding also meant that we brought to this project a realist, naturalist and cosmopolitan epistemology. Added to this, we are an interdisciplinary research team and some of our members were particularly sensitive to differences between what people say they do and what they actually do. It quickly became clear at this stage that we wanted to try to attempt at least some degree of objective measurement in our research design in order to explore what Kristján

Kristjánsson (2010) describes as the self / character realism – the view that who we are (morally) is not necessarily the same as the beliefs we have about who we are .

In the starting months of the project we also reviewed the literature on character or moral education in general (Walker et al., 2013) and the existing measures of moral judgement, character or virtue, in particular. The main options for the measurement of virtue or character that were considered and our reactions to those possibilities are shown in some detail in Table 1 below. One important consideration depicted in the table is the idea that the Jubilee Centre might develop its own test of implicit moral virtue. The aim here would have been to find a measure of moral virtue that could track emotional reactions directly - without the medium of biased rationalisations. Such an implicit measure would hone in on unconscious processes. This was an exciting prospect and would have made for a very different presentation here. However, in the end, this option was discarded as too ambitious along with some other reasons shown in Table 1².

During this review process we formed the view, in line with David Funder’s work on personality (Funder, 1995 ; Funder, 2012), that character or virtue is best measured using a multi-criterial method. Not only was this the best available approach but perhaps, also, this is the only way to attempt the measurement of something as complicated as virtue or character. One method alone may not suffice. Methodological triangulation has been around for a very long time (Campbell and Fiske, 1959 ; Webb and Campbell, 1966) but reasons for its use are varied. For example, Mathison (1988) points out that triangulation has often been employed to correct flaws or bias in component methods but advocates instead using multiple methods to deliberately generate convergence, inconsistency and contradiction of results. The aim she suggests is to provide ‘more and better evidence from which researchers can *construct meaningful propositions* about the social world’ (Mathison, 1988:15). That methodological triangulation can lead to one or more of these three

² The Jubilee Centre for Character and Values piloted a School Virtue Measure (SVM) recently with 2300 pupils.

possible outcomes (convergence, inconsistency and contradiction) is clear, but our aspiration is to measure objective virtue.

Table 1: Character Education Project - Measuring character and virtue among Year 10 (aged 15 and 16) in UK Secondary Schools

Measures considered during research design process						
Author	Institute	Measure/type	Format	Description / aim	Pros	Cons
<u>Selected</u>						
Professor Steve Thoma	University of Alabama - Office for Ethical Development	Adolescent Intermediate Concept Measure (Thoma et al., 2013) http://www.ethicaldevelopment.ua.edu/adolescent-icm	Seven moral dilemmas about social issues. Participant must rate options for the protagonists' actions and reasons for acting.	Based in the DIT family, this measure targets moral judgement and is intended to activate intermediate moral schemas.	The level of detail and range of choice options introduce real and difficult dilemmas, reflective of life. Possible outcome measure for character education programmes.	A lot of reading for participants.
Peterson, Seligman, Park	VIA Institute on Character	96-item VIA Inventory of Strengths for Youth (VIA Youth Survey) (Peterson and Seligman, 2004 ; Park and Peterson, 2006)	96 questions – self-report	Measures 24 character strengths	Off-the-shelf Universal across cultures	198 question version was too long but this shortened version was introduced during research design period.

Considered but not selected

	Jubilee Centre	New test of implicit moral virtue developed by Jubilee Centre	<p>This would be a self-report, but measuring features of character of which the respondent may be unaware. Of course participants would sign up to this.</p> <p>Latency-based measures assume cognitive priming's can reveal truths about who we really are deep down - below conscious thought.</p>	<p>For Aristotelians paradigmatic moral virtues are emotions or incorporate emotions. The aim here would be to find a measure of moral virtue that could track emotional reactions directly - without the medium of biased rationalisations. Such an implicit measure would hone in on unconscious processes.</p> <p>E.g. Implicit Association Test -or Sequential Association Test - measures how quickly respondents recognise pros and negative connotations of a stimulus adjective.</p>	<p>Risk versus gain?</p> <p>Would this target character more accurately?</p>	<p>Takes too long to develop and currently insufficient internal Jubilee Centre expertise.</p> <p>Ethical considerations: -Opt-out consent from pupils may be inappropriate? -Are such measures deceptive?</p> <p>Context of test can influence results - therefore is the test really measuring an underpinning deep reality.</p>
	The self-importance of Moral Identity 2002 (Aquino and Reed, 2002)	Social Cognitive conceptions of moral identity	<p>Participants read text about a person and are asked to show the degree to which they want to emulate the traits of the moral person. Nine characteristics are covered. Individuals then answer questions intended to assess how much the participant attempts to emulate this person.</p>	<p>Measures the degree to which moral identity is central to a person.</p>		<p>Situational – (i.e. prior events) are thought to influence results.</p>

Richard Lerner	Institute for Applied Research in Youth Development at Tufts University	Measure of Positive Youth Development - Student Questionnaire Short Version (Lerner et al., 2005)	Self-report questionnaire for pre/post intervention.	Developmental Approach Measures 5c's: competence, connection, confidence, caring, character	Available by request No training required to use / analyse Appropriate for 10 yrs and up, Good questions, easy format, well known/tested.	Little bit off-the-mark for what we want to measure. Designed for pre and post-test. Measuring more than character Cognitive focus
Rest, Narvaez, Bebeau, & Thoma	University of Alabama - Office for Ethical Development	Defining issues Test 1 and 2	This test comprises moral dilemmas each followed by 12 items (DIT 2) - participant takes role of protagonist - rate and rank items.	The Defining Issues Test (DIT) measures moral judgement that is mostly non-verbal and intuitive. It activates moral schemas and assesses them in terms of importance judgments. Participants must rate and rank items in terms of their moral importance.	Measures non-verbal and intuitive level. DIT is supposed to measure understanding at the level that drives most decisions for most people	Not suitable for children in our study.
Prof Lindt	University of Konstanz	Moral Judgement Tests http://www.uni-konstanz.de/ag-moral/mut/mjt-engl.htm#DIT	Rooted in Kohlberg tradition, this is not virtue based but is supposed to take better account of virtue ethical objections to Kohlberg than the DIT does. Comprises 2 dilemmas and 24 items	A moral competence test which assesses 2 aspects of moral judgement – cognitive and affective - at the same time. Proficiency in dual aspect	Free to use Good for 11 - 14 year olds	Experimental psychology experience necessary. Age range only up to 14.

			(arguments) to be rated by the participant.	theory of moral behaviour and in experimental psychology is necessary to use and interpret this measure.		
Gibbs and Basinger	Ohio, USA	Sociomoral Reflection Measure-Short Form (SRM-SF) (Gibbs et al., 1992)	Paper and pencil production style measure that assesses maturity of sociomoral reflection. Does not involve moral dilemmas but instead uses lead-in statements: e.g. 'let's say a friend of yours needs help and may even die and you're the only person who can save him or her'. This is followed by evaluation questions.	Assesses developmental stages of moral judgement. Participants' maturity level is measured by scoring their justifications for moral behaviours.	May be richer and more ecologically valid than the context provided by prepared dilemmas. Format seems useful.	Focuses on moral reasoning alone. The questions are a little simplistic and more appropriate for younger subjects.
Howard Curzer	Philosophy Texas Tech University	Sphere-Specific Moral Reasoning and Theory Survey (SMARTS) http://www.ttuethicsproject.com/demo.php	Twenty 2 part questions about various scenarios. Takes 30 minutes to complete. First part of each question asks if an act is right or wrong. Participant then picks why.	Measures moral theory choice and moral reasoning independently.	Close to what we are trying to do.	Probably not suitable for target age range.
Dr Robert Cloninger	Department of Psychiatry Washington University in St. Louis	Temperament and Character Inventory (TCI) http://psychobiology.wustl.edu/index.php?option=com_content&view=article&id=64&Itemid=84	Comprises tests designed to identify the intensity of and relationships between the seven basic personality dimensions of Temperament and Character, which interact to create the unique personality of an individual.	The four measured Temperament dimensions are Novelty Seeking, Harm Avoidance, Reward Dependence, and Persistence.	Supposed to bring together emotions and character?	Not clear it can be used for children. Wellbeing, happiness and therapy focussed.

Aristotle holds that there is moral truth in the degree to which a person is virtuous or not. As a sociologist, I understand this to mean that in all societies there are ways of living that will lead to a better (more flourishing) life for all concerned. These ways of living are ordered around the virtues that need to be cultivated in the young to include a period of habituation towards the development of cognition, feeling, perception and behaviour. The best case outcome will be a critical, independent and moral person capable of exercising practical moral wisdom even when this is contrary to their immediate community. We want to explore, as part of a much broader remit, some of the likely differences between what participants report about their character strengths and the extent to which they are becoming virtuous.

Triangulation has been an approach of choice in the paradigm of realism because it assumes a single reality. In the next section, I will introduce each of the three methods we triangulate including, where appropriate, a discussion of their underpinning theoretical framework in relation to our stated philosophical position. Before doing so, it is important to point out that in the original bid for funding from the John Templeton Foundation there was a stated intention to survey 10,000 pupils, visit 100 schools and interview 300 teachers. Together with our aspiration for objective measurement and methodological triangulation, this promised a research project that was going to be practically demanding not least in terms of locating, recruiting and visiting 100 schools across UK, but also in reaching such large numbers of pupils. It is also worth pointing out that we have recently learnt that the experienced Taiwanese character educationist Angela Chi-Ming Lee has opted for a similar triangulation method when designing a pilot study of the moral competence of Taiwanese students. Her instrument combines self-reports, responses to moral dilemmas, interviews and ratings by teachers (Lee, 2013). Lee's pilot study found self-report scores to be higher than dilemma scores and, arguably, over-estimating students' competence. No statistical difference was found, however, between dilemma scores and interview scores. Lee's findings are in line with the assumptions that underwrite the current research work in the Jubilee Centre for Character and Values.

Moral Dilemmas – Intermediate Concept Measure (Adolescents)

Long associated with Lawrence Kohlberg (Kohlberg, 1958 ; Kohlberg, 1969), moral dilemmas have been criticised for too much emphasis on moral judgement to the detriment of predicting moral action. The problem with the Kohlbergian use of moral dilemmas in qualitative interviews is that they only ever established weak connections between the two domains (Blasi, 1980). Neo-Kohlbergians such as Darcia Narvaez, Mickey Bebeau, Steve Thoma and the late Jim Rest have responded to these and other complaints by placing less stress on moral stages than Kohlberg did and more emphasis on context and schemas as patterns of moral functioning (see for example Narvaez, 2005). In this neo-Kohlbergian model, schemas normally remain firmly within the domain of moral judgement, but it seems still possible to our Aristotelian viewpoint that the activation of a schema in response to a moral dilemma might also involve a fuller reaction than is claimed by many Neo-Kohlbergians. Indeed, Lapsley and Narvaez tend towards this broader view of schemas that are ‘chronically accessible’ for the person pending stimulation (Lapsley and Narvaez, 2006:268). This kind of wider interpretation of schema theory fits well with our Aristotelian perspective and partially motivates our attraction to moral dilemmas based on this theory.

Moral dilemmas, such as the Defining Issue Test for professionals, have been designed to provide sufficient contextual detail in order to stimulate a person’s schema so that their responses will also include the activation of additional knowledge from the context of a participant’s life which, too, will shape what is perceived of the dilemma. The so-called Four Component Model is an important feature of this ongoing work emphasising four psychological processes of the moral person: ethical judgement, reasoning, sensitivity and action (Narvaez and Rest, 1995). The understanding underpinning the neo-Kohlbergian approach is that individuals have a general and dominant moral orientation towards their lives known as bedrock schemas. Bedrock schemas can take the form of three different types, starting with the most morally basic level of ‘personal interests’ where moral judgments are made largely on the basis of self-interest. The next type of ‘maintaining norms’

represent moral improvement but still a person with a 'maintaining norms' bedrock schema will tend towards moral judgement that concord with norms surrounding them even when those norms are morally suspect. The third and final type of bedrock schema is the 'post-conventional' level of moral judgement that is exhibited by individuals who are able to make their own considered moral judgements and, as such, are thought morally more sophisticated than those operating from the other two bedrock schemas. Neo-Kohlbergians' long association with ethical behaviour and judgement in the professions leads them also to emphasise very specific codes such as professional codes that influence a person's moral judgement in addition to their broadly intuitive bedrock schemas. In-between intuitive bedrock schemas and specific contextual codes are so called Intermediate Concepts. Neo-Kohlbergians believe Intermediate Concepts (e.g. honesty, courage or self-discipline) to be more specific to daily life than both specific codes and general bedrock schemas - and these may overlap similar virtue based concepts (Thoma, Derryberry et al., 2013). Again, this is something that piques our Aristotelian interest even though virtues from this perspective are more like concepts and as such are quite different to Aristotelian virtues that are more moral dispositions that include - in neo-Kohlbergian language – all four components from their model. For Steve Thoma, individuals interpret virtue in more specific and targeted ways.

In our research design phase, the new Intermediate Concept Measure for Adolescent populations (Thoma, Derryberry et al., 2013) appealed to us as a kind of objective measurement of the age group of interest. In its complete form, this measure comprises seven moral dilemmas which we have chosen to reduce and adapt under the helpful guidance of its authors. Out of seven possible dilemmas, we have used only three for two key reasons. First, school timetabling meant that the combined surveys (UK ICM (Ado) and VIA Youth Survey) could not last for more than forty minutes. Second, although we were generally optimistic about pupils' capacity to concentrate for lengthy periods because they regularly do so for exams etc., we judged that three moral dilemmas plus the VIA Youth Survey would be on the edge of possibility for less able pupils or for those pupils who find reading difficult or unpleasant. Each of the chosen moral dilemmas targets a specific intermediate

concept or virtue: self-discipline, courage and honesty. These were selected because they clearly matched qualities measured by the VIA Youth Survey, seemed relatively uncontroversial, and were associated with stories that were easily relatable to UK Year 10 pupils. Aside from dropping four of the original dilemmas, we also replaced uniquely American terms with British ones and swapped the courage dilemma about a part in a play for one about a female gymnast who feels her coach is exploiting the girls' good looks for publicity. We feel this communicates a more concrete situation. Notwithstanding these changes, the structure of the original measure has been retained and Professor Thoma has approved our amendments. The final version of this shortened UK ICM (Ado) is in electronic form, including an audio version for those that need it. This survey has also been administered personally by our researchers who visit each school, introduce the measure and are present throughout to answer questions of clarification and to be sure that participants complete the survey on their own. As for the original version, participants must read each dilemma and rate action choices and reasons for action choices from 1 (I strongly believe that this is a GOOD choice/reason) to 5 (I strongly believe that this is a BAD choice/reason). They must then select and rank best (first, second and third) and worst (and second worst) options for actions and reasons. During analysis, responses are compared with previously established expert panel decisions about responses that are adequate, inadequate or neutral. These are then scored. Best and worst scores for choices and reasons are used to achieve a total good and total bad score that determines the extent to which answers correspond or contrast with the expert panel. Importantly, there is no overall right or wrong answer to the dilemmas. A total ICM score is also calculated.

We are hopeful that if Narvaez's broad interpretation of schemas is sustainable, then this measure might also be taken to evoke grand moral schemas, covering more or less the whole of moral functioning. Virtues for Narvaez are not secondary to bedrock schemas or narrow; but are 'the moral habits of virtue theory and are social cognitive schemas whose chronic accessibility favors automatic activation' (Lapsley and Narvaez, 2006:268). In other words, moral schemas might then

be basically the same as Aristotelian *hexeis* – or, if you are a sociologist, Bourdieuean *habitus*, although of course this remains to be seen or more importantly demonstrated.

Self-Reporting – Values in Action Inventory for Youth (VIA Youth Survey)

We also want to hear from the pupils about their own characters. The VIA classification is well tried and tested and is ‘grounded in a long philosophical tradition’ (Peterson and Seligman, 2004:9) going back to Aristotle. Positive psychologists have accumulated a substantial database of self-evaluations of moral virtue from all over the world, which make for exciting comparisons and correlational studies – we see this as a significant strength of using established measures such as this. Honing in on six broad categories of virtue, the VIA measures in general target twenty-four character strengths that are processes making up the virtues and which cover specifically the areas we target in the moral dilemmas (courage, honesty, self-discipline), but also provide a holistic self-assessment. Work is currently underway to undertake a substantial revision of the VIA-IS over the next year and Bob McGrath who will undertake this effort also has papers in press concerning VIA measures (McGrath, in press ; McGrath, in press). One challenge associated with triangulating our methods has been to contain the overall fieldwork to a level of time and effort that is feasible for both schools and pupils of all abilities. Having reduced ICM (Ado) to three dilemmas, we were grateful that the short version of the VIA Youth Survey became available just as we needed it. Eventually, with kind permission from Dr Ryan Niemiec and Kelly Aluise from VIA, we have incorporated this into our electronic survey so that the pupils move seamlessly from ICM UK (Ado) to the 96 short questions of the VIA Youth Survey, answerable by choosing from five options to show whether various statements (e.g. ‘I don't boast about what I achieve’) are ‘very much like me’ through to ‘not like me at all’. This makes for a helpful change of pace from the moral dilemmas, especially for some pupils who find reading difficult or uninteresting.

Teacher Interviews

Finally, three teachers of the pupils surveyed at each school are interviewed using a mixture of open and multiple choice questions. The interview is designed to cover a lot of ground in a short time period which is frequently necessary to avoid disrupting teaching. Teachers are shown interview response cards containing multiple choice questions. When time is available the interview is expanded to gather more detailed responses. In addition to asking about teachers efforts to develop character, we also ask them to select from Peterson and Seligman's list of twenty-four character strengths the three most and three least pronounced strengths among the group of Year 10 pupils that they are teaching. Further, we ask of the teachers how often certain qualities have been displayed by the same group of pupils over the last few months. The qualities asked about are:

Optimism about the future

Empathy for others in the classroom

Good understanding of their own feelings

Eager for new knowledge

Honesty

Modesty and humility

Self-control

Delayed gratification

Gratitude

Importantly, we do not ask teachers to comment on individual young people, because we would have found this ethically questionable and, similarly we suspect, the teachers would have found this uncomfortable or undoable.

Cases

Identifying and recruiting schools is an ongoing effort. Our aim is to visit fifty secondary schools as half of our overall sample. Even so, I agree with Nick Emmel (2013) that 'it is not the number of cases that matters, it is the work they are shown to do in interpretation and explanation that

counts'. Of more importance, then, is the final mix of schools in terms of the kinds of educational experiences that they represent in relation to experiences possible or likely for Year 10 pupils across UK. Anyone involved with researching UK schools will know that engaging such busy institutions can be difficult, especially involving surveys that require computer and internet facilities together with forty minutes of teaching time.

As well as trying roughly to match our school choices to the population (e.g. ensuring good representation in London etc.), we are actively seeking out variation between schools in terms of specific criteria such as, for example: state school versus independent, faith-based schools of different types, grammar schools, single-sex versus co-educational, rural versus city, affluent versus deprived areas, and so on. In order to push and stretch the boundaries of the experiences of the Year 10 pupils and their teachers available for our research, we are using "purposive" case selection to gather rather routine and unusual (Becker, 1998) cases that are to be actively sought if they are considered lacking from our overall choice of schools. For example, we actively seek the inclusion of schools that adopt uncommon philosophical approaches such as, for example, a Steiner school.

Regardless of our criteria for selecting cases, we cannot influence the reasons why schools will or will not take part. Some schools, especially those that share our emphasis on the flourishing child, are only too keen to become involved. Other schools, however, are much harder to reach - either because they are genuinely too busy doing their regular jobs or for practical reasons such as, for example, our researchers cannot get past vigilant gatekeepers who are used to numerous enquiries each week from would-be researchers and other interested parties. All of this makes for a mixed bag of approaches to the schools. For every school that we have so far visited, we have had significant conversations with six others that didn't materialise into definite research visits. Moreover, hundreds of unsolicited letters yielded no school visits. The most successful recruitment technique has been all different types of personal contact and snowball sampling from schools previously encountered. Other networks have been used too, including accessing schools via

teachers known to friends and colleagues of researchers or through a number of official networks such as for example, Birmingham University Alumni, Religious Education, maths and science networks, Academy school groups and educational / teacher literature.

RESULTS

In this part of the paper, I shall focus on seven hundred and fifty pupils from eight schools, together with thirty-two teachers from those schools as a sub-sample of the total data available (total numbers are given in the introduction). For the purposes of this conference I am focussing only on data that has been very recently generated in order to provide a snapshot of all three parts of our research design (moral dilemmas, self-reporting and teacher interviews) at this mid-point in our data collection period. This is because these data have been gathered using improved methods, especially the incorporation of the 96-item VIA Youth survey into our own survey instead of asking pupils to access the VIA website directly during our research visits to schools. Less recent data that has been gathered through the VIA website requires additional preparatory work. We are grateful to the VIA Institute on Character for making this improvement possible by introducing a previously unavailable process of bulk-data entry³.

Of the eight schools represented in this recent data, three are prestigious independent (fee paying) schools in the north and south of England and five are large state secondary schools - three in different parts of England and two in different parts of Scotland. Two of the English secondary schools are catholic and one is a business and enterprise college. One of the Scottish schools is an academy and the other a large high school.

³ We are especially grateful to Kelly Aluise and Eric Zimmerman for their ongoing support, guidance and help.

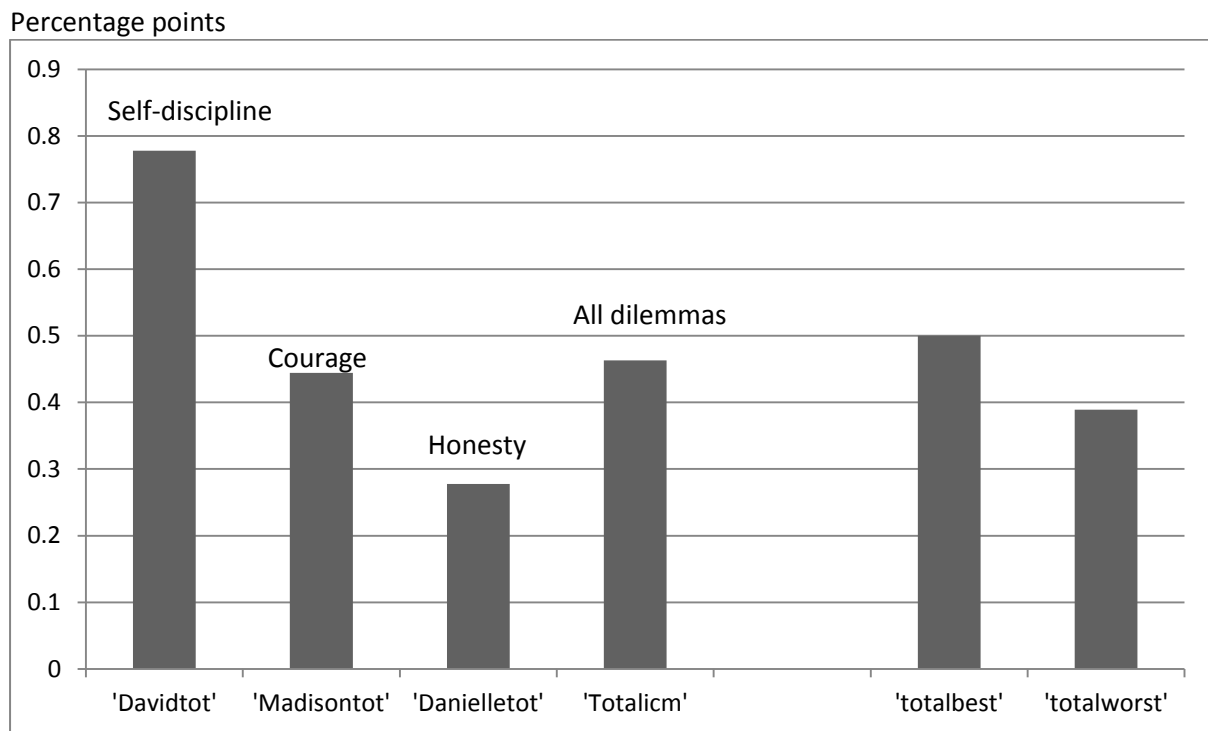
Overall Comparisons

UK ICM (Ado)

The first dilemma, measuring self-discipline, concerns for David a decision between studying for an imminent exam versus accompanying his girlfriend on an outing. The second dilemma, measuring courage, involves for Madison a decision about if she should speak out against her gymnastic coach who is exploiting only the girls' looks for publicity. The final dilemma, measuring honesty, concerns for Danielle questions associated with cheating in an exam and what she should do about knowledge that this has occurred. Total median scores have been calculated for each dilemma. These include responses to both choices and justifications and require the selection of both good and bad options. Total scores for each of the dilemmas are labelled in subsequent charts as: 'Davidtot', 'Madisontot' and 'Danielletot'. As can be seen from Chart 1 below, respondents have made judgements that are closest to the expert panel judgements for the dilemmas measuring self-discipline (0.78), then for courage (0.44) and finally for the dilemma measuring honesty (0.28). One hundred percent agreement with the expert panel would produce a score of 1.0.

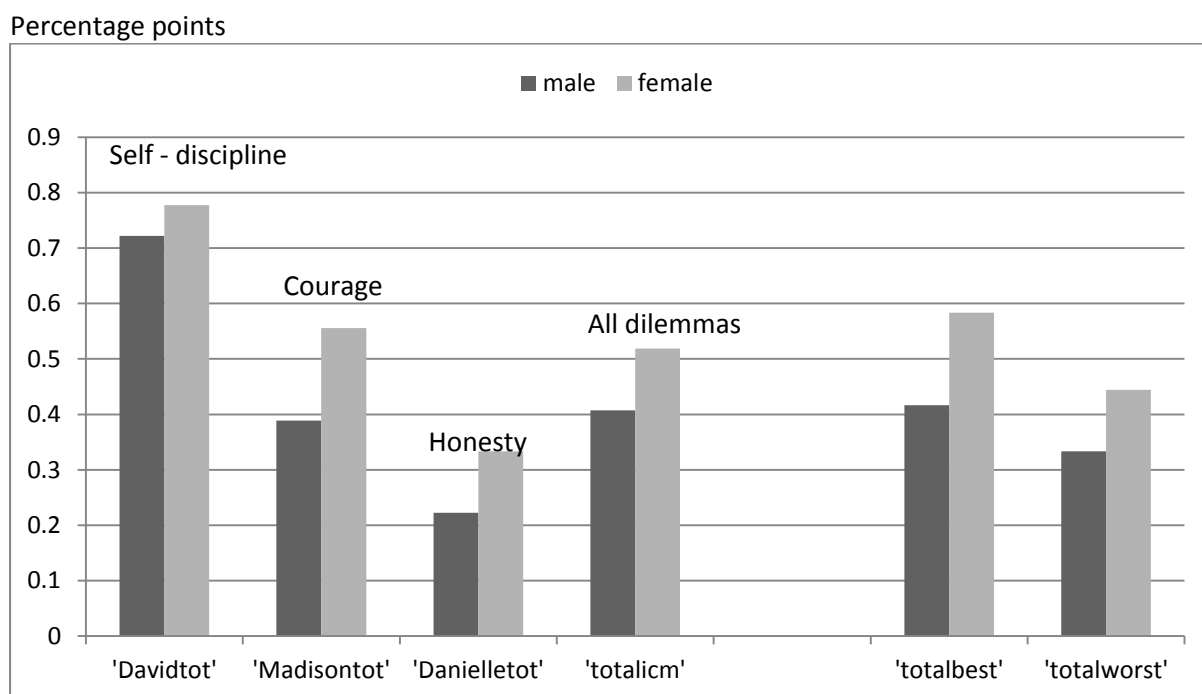
As can be seen from Chart 1 below, differences in total responses to the three moral dilemmas are clear. Honesty scores are very low and self-discipline scores are very high. Overall, the 'totalicm' score, depicting the level of choice and justification judgements (both good and bad) across all of the dilemmas is 0.46 which means that compared to the expert panel, individuals are on average appropriately identifying both acceptable and unacceptable items at a rate of 46%. Once median scores for best and worst choices and justifications have been separated out, the pupils' identification of best choices matches more closely expert judgements than does the identification of worst choices.

Chart 1 – Total median scores for each dilemma; across all dilemmas; and separated best choices and separated worst choices across all dilemmas



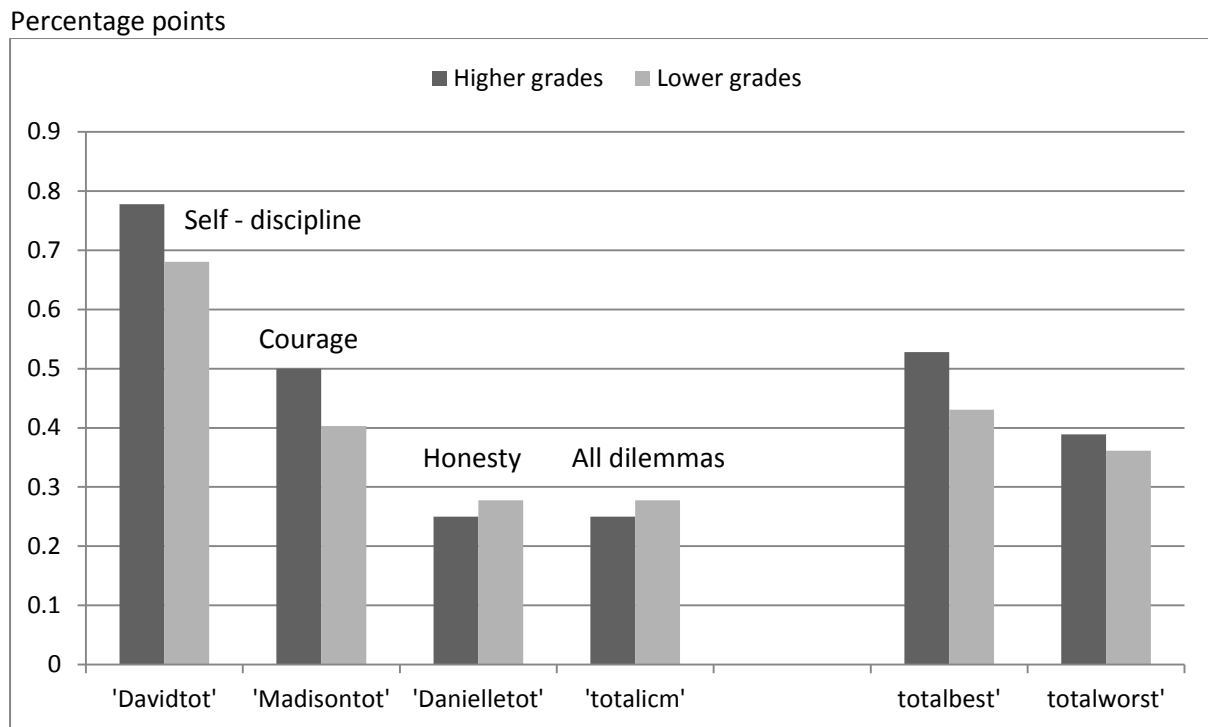
For all of the total scores so far discussed, female pupils selected choices that were closer to the expert panel than did male pupils (see Chart 2 below). This was especially noticeable for Madison’s dilemma intended to measure courage. It is worth noting that this dilemma is especially pertinent to females because it centres on a gymnastics coach using the good looks of the girls for publicity purposes. In contrast, the smallest gender difference has been recorded for David’s moral dilemma measuring self-discipline.

Chart 2 – Total median scores for each dilemma; across all dilemmas; and separated best choices and separated worst choices across all dilemmas by gender



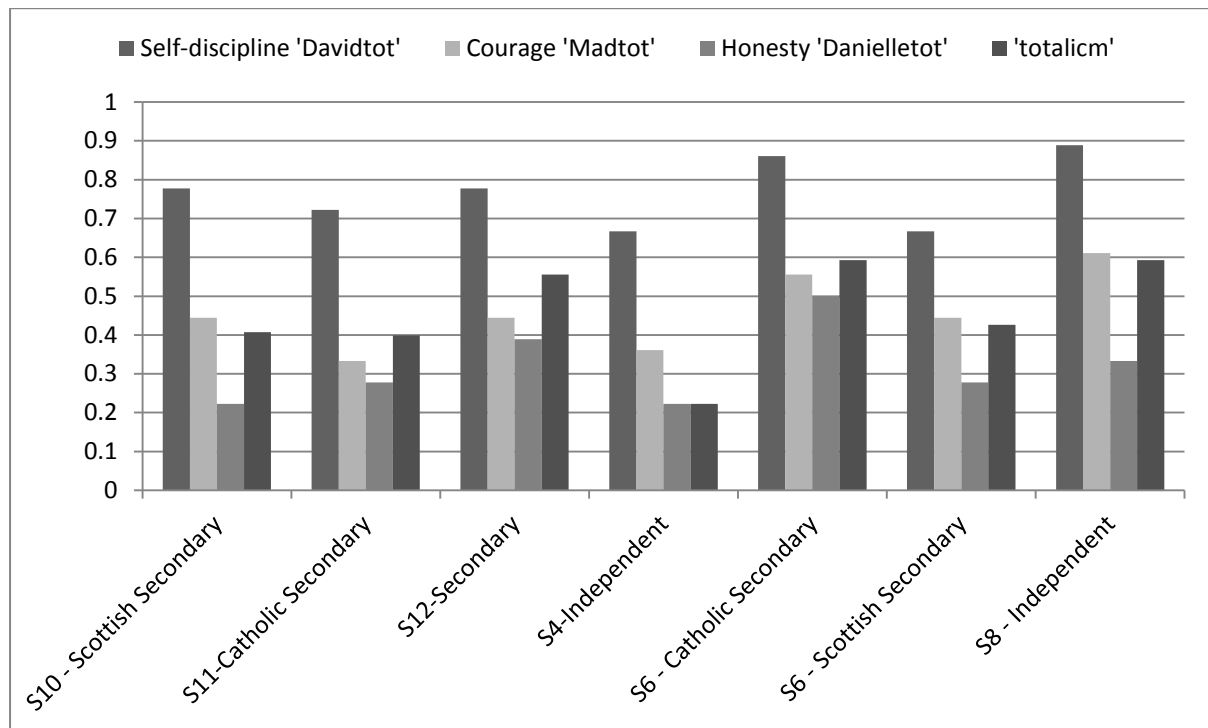
All participants were asked in a demographic section of the survey if their school grades when compared to other pupils were better; mostly better; about the same; or somewhat lower. After recoding these options into two new variables (higher grades and lower grades) and calculating median scores, those pupils who said they were getting grades that were better or mostly better than others (higher grades) also achieved median scores that were closer to the expert panel for self-discipline and courage (by almost 10 percentage points) than those that said they were getting grades that were about the same or somewhat lower than others (lower grades). However, those claiming higher grades scored slightly less than the lower grades' group for the dilemma measuring honesty as well as for the total of all of the dilemmas ('totalicm'). Caution is needed interpreting these results given their partial and tentative status. In time, we will want to check the worrying possibility that pupils with better grades are more able to work out adequate answers but are not necessarily capable of better moral judgement.

Chart 3 – Total median scores for each dilemma; across all dilemmas; and separated best choices and separated worst choices across all dilemmas by school grades



Comparing total median scores for responses to each moral dilemma by school doesn't alter the general pattern towards noticeably higher scores for self-discipline and lower for honesty with courage in-between (Chart 4 below). One of the independent schools is represented by only a small sample of boys and has been excluded from this particular comparison (though this pattern of dilemma scoring does show there too). It is interesting to look at 'totalicm' scores by school that are above and below the median for the whole sample (0.46) because this does not seem to relate to types of school in any stereotypical or expected way. For example, although one independent school scores highly, the other does not. Similarly, one catholic school scores highly while the other does not. Both of these types of school are generally believed to be more concerned with the development of character. Scores between Scotland and England are not noticeably different either. There are too few schools to make much of this data at this stage but it is hoped that the combination of a lot of different types of schools will eventually show trends and patterns that can be connected to what is different about those schools.

Chart 4 – Total median scores for each dilemma; across all dilemmas; and separated best choices and separated worst choices across all dilemmas by school



Overall, the UK ICM (Ado) appears to be working much like the full USA version. For example, there are trends among this sample that match younger USA high school students such as females scoring higher than males and the tendency towards better recognition of good choices or justifications over bad ones. These are encouraging signs for this reduced and slightly adapted measure.

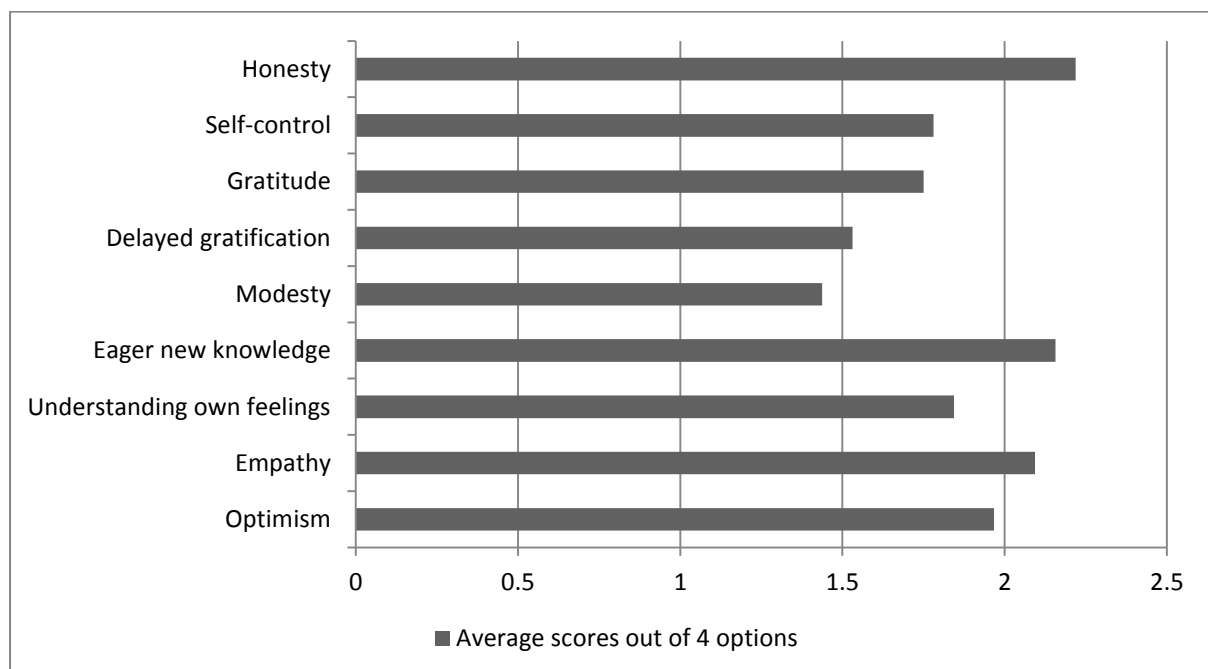
Teacher interviews

Thirty-two teachers were asked directly about the same Year 10 pupils surveyed by the UK ICM (Ado) measure and the VIA Survey for Youth in two different ways. First, teachers were asked how often in the past few months (before the day of the surveys and interviews) had certain qualities been displayed by the pupils. A detailed list of the qualities asked about is provided in an earlier section of this paper about teachers. Second, the teachers were asked to select from Peterson and Seligman’s list of twenty-four character strengths the three most and the three least pronounced strengths among the group of Year 10 pupils they were teaching.

Qualities recently noticed among pupils by teachers

Average (mean) scores were calculated for all responses from the thirty two teachers for each personal quality noticed among the pupils they were teaching. Complete results are shown in Chart 5 below. Teachers could choose from four options (1 (a lot) to 4 (very little)). Scores have been reversed so that a higher number corresponds to higher average reports of the named quality. In contrast to low UK ICM (Ado) scores for this virtue, honesty is rated as highly present among the pupils by their teachers in the past few months. Moreover, self-control (self-discipline) is rated by the teachers as relatively lacking among the pupils. Compared to the pupils' UK ICM (Ado) scores this represents a reversal of emphasis: pupils score low for honesty and high for self-discipline for the moral dilemmas but their teachers say that they have noticed more honesty than self-discipline among the pupils over the previous months. Teachers were not asked about courage in this particular question so it cannot be directly compared here. (They are, however, asked if bravery is prominent or not from the list of twenty-four strengths below).

Chart 5 – Personal qualities noticed among pupils by their teachers in the past few months



Teacher reports of pupils' most and least prominent character strengths

More than anything else the teachers report that from the list of twenty-four character strengths, the following were most prominent:

humour, fairness, creativity, curiosity and open-mindedness.

They also report as least prominent:

persistence, spirituality, self-regulation, gratitude, open-mindedness and leadership.

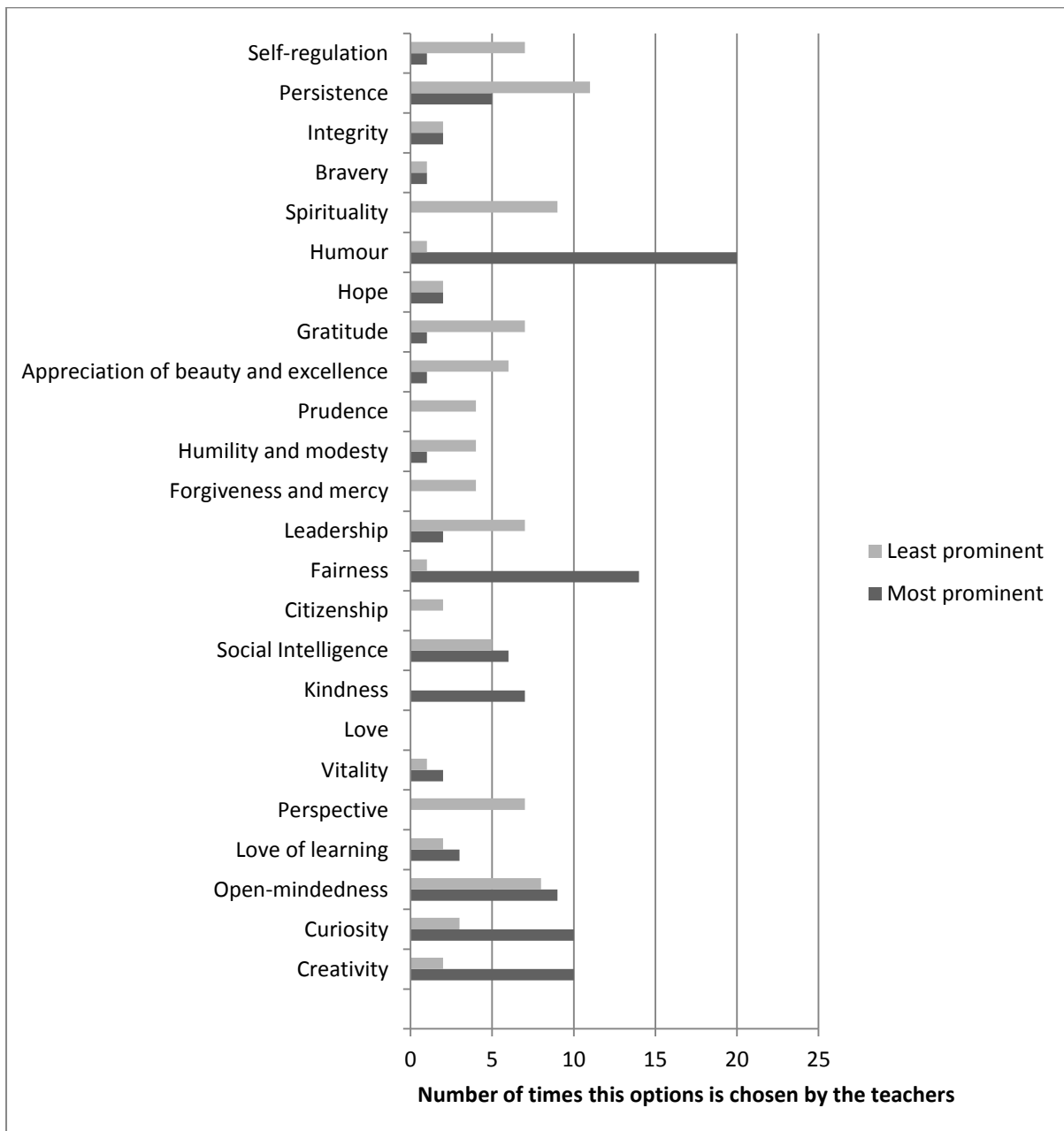
It is important to remember that the teachers were asked to scan the complete list of twenty-four character strengths and can select only six: three that they believed were prominent and three that they thought were not among the Year 10 pupils that they were teaching. Low scores (see Chart 6 for complete results) suggest that the strength in question has barely been selected and so is either not something of relevance or notice for the teachers or is not especially present or absent among the pupils. In terms of comparisons with the results from the UK ICM (Ado), we are particularly interested in the virtues of self-regulation (self-discipline), bravery (courage) and honesty⁴.

Persistence⁵ may also be connected with self-discipline. Both persistence and self-regulation were reportedly not prominent among the pupils, whereas bravery and honesty (integrity) were barely selected out, perhaps having been trumped by other options given that the teachers were asked only to select six character strengths in total.

⁴ The term shown in brackets is the name given to the virtue measured in the UK ICM (Ado), whereas the preceding term is used in the list of the twenty-four character strengths and the VIA Youth survey results.

⁵ In the results returned from the VIA, persistence is referred to as perseverance and we are using the terms interchangeably.

Chart 6 – Reports by teachers of pupils’ most and least prominent character strengths



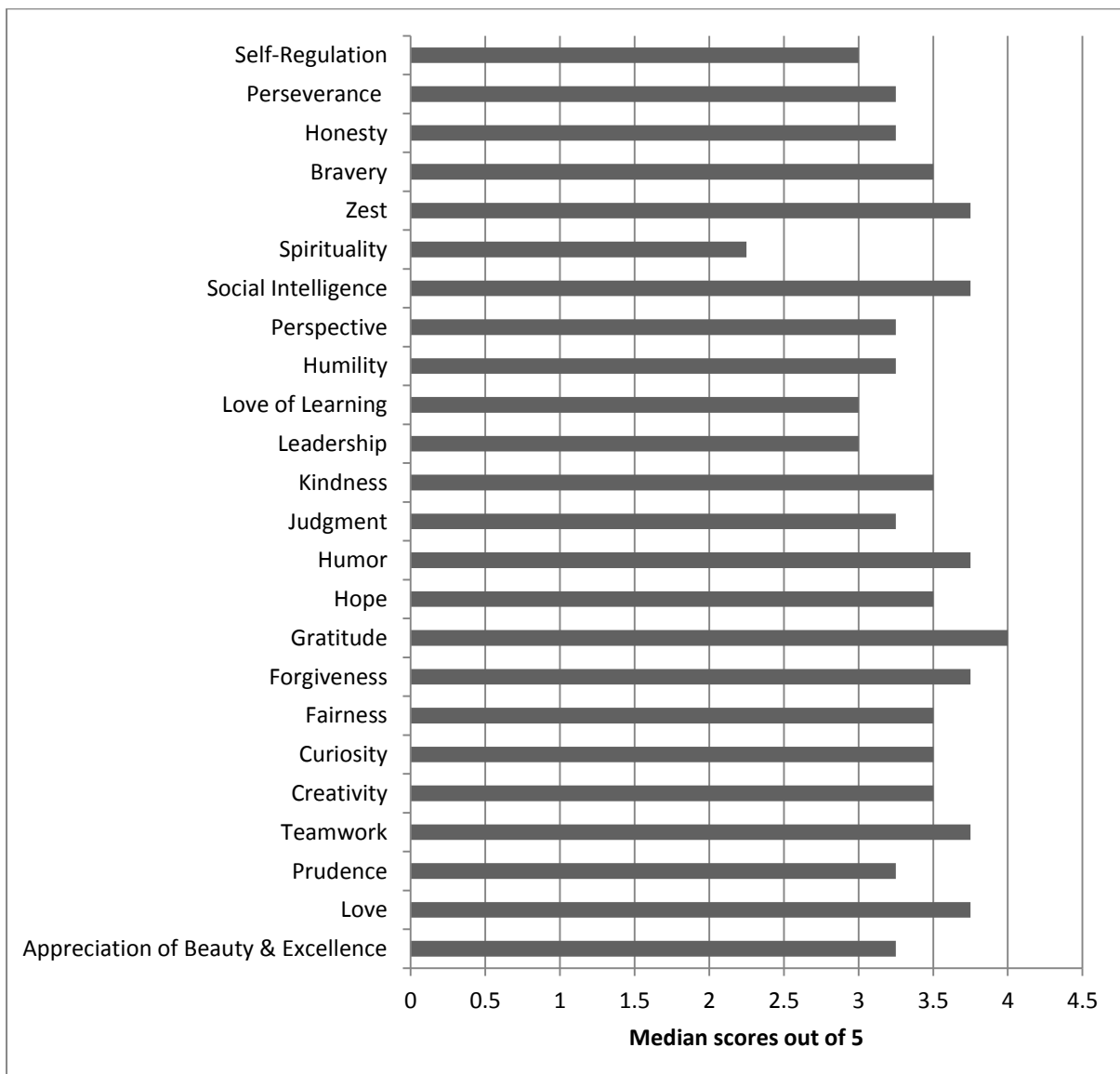
96-Item VIA Youth Survey

Somewhat expectedly, pupil’s collective reports of their own character strengths overall do not show extreme deficits, although spirituality has by far the lowest median score (see Chart 7 below for complete results). The pupils rate themselves most as having the following strengths:

Gratitude, humour, zest, teamwork, love, social intelligence, forgiveness.

How did the pupils rate themselves in terms of the three virtues measured in the UK ICM (Ado) - honesty, courage, self-discipline? These results, along with results for perseverance, are shown at the top of Chart 7 below. Although bravery is rated more highly than honesty which is rated above self-regulation, the differences are small and perhaps of more relevance is that none of the three character strengths (or virtues) stand out particularly for better or worse from all of the others.

Chart 7 – Pupils’ reports of their own character strengths – 96-item VIA Youth Survey



Now that brief and tentative results have been set-out from all three research methods, it will be helpful to summarise these ahead of a discussion about how each might relate to the other in the next section. This summary is provided in Table 2 below.

Table 2 – Combined summary of results from all methods for honesty, courage and self-discipline

UK ICM (Ado)	honesty	courage	self-discipline
Median scores	0.28	0.44	0.78
Interpretation	low	average	high
Teacher Reports			
-Recently noticed pupil qualities (out of 9)	honesty	not asked	self-control
Mean scores	2.22	.	1.78
-Most / least prominent (out of 24 character strengths)	integrity	bravery	self-regulation (SR) & perseverance (P)
Count most, least	2,2	1,1	SR: 1,7 P: 5, 11
Interpretation(combined)	rated highly when asked directly	hardly selected	SR: lacking P: highly lacking
VIA Youth Survey			
Median scores	honesty 3.25	bravery 3.5	self-regulation 3
Interpretation	average	slightly above average	slightly below average

DISCUSSION

Overall, the scores for UK ICM (Ado) are possibly the most interesting single result so far, especially large differences between total results for each dilemma. This is the only measure used in our triangulation research design that is objective, and that is intended to measure or correspond to action. For example, the long line of DIT tests on which this measure is built, have produced scores that correspond to behavioural and decision-making variables (Rest, 1986 ; Thoma, 2006).

Moreover, in Professor Thoma's paper for this conference, he claims that there is a relationship

between ICM scores and behaviour or action. This is argued on the basis that students who had been placed in school suspension – as indicative of a tendency towards making poor choices – also had difficulty selecting out good and poor choices in the ICM. This is used by Professor Thoma as some evidence that students who objectively make poor choices do so also on the ICM measure. It is further noteworthy that high scores for self-discipline and low for honesty in the UK ICM (Ado) measure mirror results from the Jubilee Centre for Character and Values School Virtue (pilot) Measure (SVM) presented at the start of this conference. The SVM provides participants with shorter moral dilemmas and asks them to choose from five options, each pertaining to types of Aristotelian virtue (from self-interest through to ‘full-virtue’). These incorporate both feelings and reasons for acting and are ranked by the pupils.

If we are to accept that ICM results do in some way measure objective virtue, then these pupils are much more likely to display in their lives qualities of self-discipline than they are courage or honesty. It may be relevant that self-discipline can be separated from courage and honesty as a performance virtue rather than a moral virtue. It could also be pertinent that self-discipline or perseverance is a much emphasised personal quality in schools and for achieving good school work. In comparing the UK ICM (Ado) results for self-discipline with both kinds of teacher reports on the same, it is clear that teachers report that self-discipline - as a character strength or virtue - is somewhat lacking in their pupils. This is even more noticeable if perseverance is brought into consideration as a personal quality that is strongly aligned with self-discipline. A look at our third data type - VIA self-reports - shows in itself a rather unremarkable result for all of the virtues of current interest because they are each reported by the pupils as moderately present in their characters. However, this result seems more interesting if it is set against UK ICM (Ado) results as well as the teacher reports. Remaining focussed on self-discipline for a little longer we can see from summary Table 2 that when, as a group, the pupils report in the VIA survey that they are averagely self-regulated they are underestimating themselves in comparison with their UK ICM (Ado) scores; and the same can be said of the teachers, though seemingly to a greater degree.

I shall now turn my attention to measures relating to honesty. According to the UK ICM (Ado) this virtue prompts very low scores from the pupils but on the contrary it is recorded as strongly present among pupils in the past few months by their teachers. Even so, honesty seems to be dropped from teachers' notice or emphasis when they are asked to choose absent and present character strengths from all twenty-four options. On average, these teachers appear to have been drawn to other character strengths that trump honesty - and bravery too. Even so, we can see that honesty is rated by the pupils as averagely present in their own characters – a result that now seems like a possible overestimation according to their UK ICM (Ado) scores for the same virtue.

Less can apparently be said about courage from an examination of these tentative and early triangulated results but it looks as if UK ICM (Ado) scores are matched more accurately by pupils' self-reports of the same⁶.

Of course, much caution is necessary when interpreting these results for a number of reasons that is discussed in the next section. Nevertheless, the purpose of this paper has been to explore and discuss at the half-way point of our data analysis, the merits and possibilities of a triangulated research design. In the exploratory spirit of this paper and conference, it appears as if the performance virtue of self-discipline may well be dominating the attention of teachers and the moral judgement – perhaps actions - of the pupils. This is what the teachers notice and value most, and it is what the pupils appear to excel in at least according the moral dilemmas. Nevertheless, these teachers report that they continue to believe it is lacking. Perhaps, important moral virtues such as honesty and courage are taking a much lower priority than an understandable emphasis on performance virtues such as self-discipline.

⁶ Teachers were not asked about courage noticed among their pupils during the past few months and barely selected it out at all above other competing character strength options.

Limitations

One possible limitation is in the moral dilemmas themselves. Might there be something in them that is causing particular difficulties or responses that are more to do with the dilemmas and their options than the respondents? This possibility will need to be explored more fully, and perhaps especially for Danielle's dilemma (honesty) which is attracting the lowest scores. Quite clearly, this dilemma does contain some options that could be viewed by the young people as 'grassing' or telling on their friends to a teacher and as such represent choices that may directly contradict strong local moral codes. This kind of nuance has not yet been explored. This is partly because this exploration of the data has been done at a high level of abstraction and, as such, does not focus on individual pupils across the measures, or on particular patterns of answering. There is much more to be done, too, in terms of exploring results between schools. As with any research project, there are aspects we could have done better and not asking teachers directly about levels of courage or bravery among their pupils in the past few months is a small oversight that is somewhat compensated by other teacher reports on most and least prominent character strengths.

CONCLUSION

In this paper I have discussed the Aristotelian inspired endeavours of (part of) the Character Education in UK Schools Project towards measuring objective virtue, together with our suggestion that methodological triangulation is currently the most promising research design to achieve this. In other research, methods such as self-reporting and moral dilemmas have been shown deficient in some regard when used in isolation and this paper has explored how two such credible and validated measures are currently being used with UK Year 10 pupils in conjunction with teacher reports on those same pupils.

Our remit for the broader research project is to explore some of the likely differences between what participants report about their character strengths and the extent to which they are becoming

virtuous. One trend explicated in this paper is the possibility that these teachers – and possibly the schools too – might be preoccupied with the performance virtue of self-discipline that corresponds with the pupils' high performance on the UK ICM (Ado)moral dilemma test (and the Jubilee Centre SVM pilot study). This emphasis may be matched by less focus by teachers on moral virtues such as honesty and courage. Even though teachers rate honesty as highly present among their pupils (in the past few months), this picture is contradicted by actual performances on the moral dilemma tests. Moreover, when faced with a choice of all twenty-four character strengths, honesty and bravery are barely chosen at all by the teachers as a most or least prominent character strength.

References:

- Aquino, K. R. and A. Reed (2002). "The self-importance of moral identity." **83**(6): 1423-1440.
- Becker, H. S. (1998). Tricks of the Trade: How to think about your research while you're doing it. Chicago, University of Chicago Press.
- Blasi, A. (1980). "Bridging moral cognition and moral action: A critical review of the literature." Psychological Bulletin **88**: 1-45.
- Campbell, D. T. and D. W. Fiske (1959). "Convergent and discriminant validation by the multitrait-multimethod matrix." Psychological Bulletin **56**(2): 81-105.
- Emmel, N. (2013). Sampling and choosing cases in qualitative research - a realist approach London, SAGE.
- Funder, D. (1995). "On the accuracy of personality judgement: A realistic approach." Psychological Review **102**(4): 652-670.
- Funder, D. (2012). "Accurate Personality Judgment." Current Directions in Psychological Science **21**.
- Gibbs, J. C., K. S. Basinger, et al. (1992). Moral maturity: Measuring the development of sociomoral reflection. Hillsdale NJ, Lawrence Erlbaum Associates, Publishers.
- Kohlberg, L. (1958). The development of modes of moral thinking and choice in the years 10 to 16. PhD, University of Chicago.

- Kohlberg, L. (1969). Stage and sequence: The cognitive development approach to socialization. Handbook of socialization theory. D. A. Goslin. Chicago, Rand McNally.
- Kristjánsson, K. (2010). The Self and Its Emotions. Cambridge, Cambridge University Press.
- Lapsley, D. K. and D. Narvaez (2006). Character Education. Handbook of Child Psychology W. Damon and R. Lerner. New York, Wiley.
- Lapsley, D. L. and D. Narvaez (2006). Character education. Handbook of child psychology: Child psychology in practice K. A. Renninger and I. E. Sigel. Hoboken, NJ, John Wiley & Sons.
- Lee, A. C.-M. (2013). A pilot study on moral thinking and communication (MTC) competence of Taiwanese students Association for Moral Education annual conference. Montreal.
- Lerner, R. M., J. V. Lerner, et al. (2005). "Positive youth development, participation in community youth development programs, and community contributions of fifth grade adolescents: Findings from the first wave of the 4-H Study of Positive Youth Development." Journal of Early Adolescence **25**(1): 17-71.
- Mathison, S. (1988). "Why Triangulate." Educational Researcher **17**(2): 13-17.
- McGrath, R. E. (in press). "Character strengths in 75 nations: An update." Journal of Positive Psychology
- McGrath, R. E. (in press). "Scale - and item - level factor analysis of the VIA Inventory of Strengths." Assessment.
- Narvaez, D. (2005). The neo-Kohlbergian tradition and beyond: Schemas, expertise, and character. Nebraska Symposium on Motivation, Vol 51: Moral Motivation through the lifespan. C. Pope-Edwards and G. Carlo. Lincoln, NE, University of Nebraska Press.
- Narvaez, D. and J. Rest (1995). The four components of acting morally. Moral behavior and moral development: An introduction W. Kurtines and J. Gewirtz. New York, McGraw-Hill.
- Park, N. and C. Peterson (2006). "Moral competence and character strengths among adolescents: The development and validation of the Values in Action Inventory of Strengths for Youth " Journal of Adolescence **29**: 891-905.
- Peterson, C. and M. E. P. Seligman (2004). Character strengths and virtues: A handbook and classification. Oxford Oxford University Press.
- Rest, J. (1986). Moral development: Advances in research and theory. New York, Praeger.
- Thoma, S. (2006). Research on Defining Issues Test. Handbook of Moral Development. M. Killen and J. G. Smetana. Mahwah, New Jersey, Erlbaum.

Thoma, S., W. P. Derryberry, et al. (2013). "Describing and testing an intermediate concept measure of adolescent thinking." European Journal of Developmental Psychology **10**(2): 239-252.

Walker, D., M. Roberts, et al. (2013). "Towards a new era of character education in theory and in practice." Educational review **Published online: 30 Sep 2013.**

Webb, E. J. and D. T. Campbell (1966). Unobtrusive Measures. Chicago, Rand McNally.