A study of moral reasoning among secondary students in a public co-educational and private girls school in Mexico

David Ian Walker & Daniel Moulin-Stożek

To cite this article: David Ian Walker & Daniel Moulin-Stożek (2019): A study of moral reasoning among secondary students in a public co-educational and private girls school in Mexico, Compare: A Journal of Comparative and International Education, DOI: 10.1080/03057925.2019.1696667

To link to this article: https://doi.org/10.1080/03057925.2019.1696667

Published online: 28 Nov 2019.

Submit your article to this journal

Article views: 24

View related articles

View Crossmark data
A study of moral reasoning among secondary students in a public co-educational and private girls school in Mexico

David Ian Walkera and Daniel Moulin-Stożekb

aEducational Psychology, University of Alabama, Tuscaloosa, USA; bFaculty of Education, University of Cambridge, Cambridge, UK

ABSTRACT

Proponents of character education claim cultivating virtues during schooling helps students, schools and society flourish but critics argue character education programs implicitly justify social inequality by assuming success or failure in life is due to individual character. There is little empirical research about which individual factors, such as gender, or contextual factors, such as school type, may affect secondary school students’ character. This article begins to address this with a comparative case study of moral reasoning within and between two different secondary schools in Mexico – a public co-educational school and a private girls’ school. Results suggest individual differences (gender, religion, family circumstances and socio-economic status) and school (moral education program and school ethos) may relate to students’ ability to make moral judgements in response to realistic moral dilemmas. We consider these findings, focusing on how gender, religion and school-type may impact on adolescents’ moral reasoning.

KEYWORDS

Character education; moral reasoning; gender; social inequality; Mexico

Introduction

Character education – the cultivation of personal attributes deemed essential for personal success and societal flourishing – has become of increased interest to educators, policymakers and researchers globally (Kirchgasler 2018). According to character educators, desired qualities to be promoted in schools can be divided into performance, civic and moral attributes (Berkowitz 2011; Seider 2012), representing a convergence of positive psychology, citizenship education and the classical aim of education as the attainment of virtue (Moulin-Stożek 2018a). Although a now burgeoning field worldwide, character education has been subject to substantial criticism. There is concern over the methodological rigour of the scientific study of character development, particularly across different cultural contexts (Clement and Bollinger 2016; Duckworth and Yeager 2015). Others have been critical of attempts to promote character in contexts of high social inequality, claiming that identified character attributes, rather than being scientific concepts, simply signify desirable behavioural norms that belong to higher socio-economic groups. Character education programs may consequently represent a form of unreflective cultural colonisation that seeks to
justify social inequity by locating responsibility for success or failure in the presence or absence of certain individual traits (Alexander 2016; Camfield 2015; Kirchglasser 2018).

Despite these concerns, little empirical research has been undertaken to establish which individual and contextual factors affect adolescents’ character – particularly in regard to moral virtues which are sometimes considered to be the most problematic kind of attribute to measure (Curren and Kotzee 2014). While most education systems have legally codified aims to cultivate moral character in some sense, reliable and objective knowledge about moral aims, related developmental processes, and measurable moral outcomes remain out of reach for schools and scholars alike. Among these identified challenges are the influence of individual differences; conflicting ideological positions about how, and if, students ought to receive moral instruction; uneven childhood development; difficulties of assessment; and patchy application of school interventions and programs (Kristjánsson 2015; Walker and Thoma 2018).

In order to begin to explore these pressing issues in a real-world context, we examine differences in moral reasoning among students in two very different kinds of secondary school in a provincial city in central Mexico (Ciudad Central): a small private girls’ school with a Catholic ethos (Colegio Privada), and a large public co-educational secular school (Secundaria Pública). We do this in order to investigate how individual differences (gender, religion, family circumstances and socio-economic status) and school (moral education program and school ethos) may be related to students’ moral reasoning performance. We are particularly interested in differences between girls’ scores in a mixed school and those in a girls’ school. To conduct the research, we employ the Intermediate Concept Measure for Adolescents (AD-ICM) as part of an embedded case study design. This instrument, already used in USA, Macedonia, Taiwan and UK, measures adolescents’ judgement of three virtues frequently emphasised by character educators: honesty, self-discipline and moral courage (Thoma and Walker 2016; Thoma et al. 2019).

ICM data had not previously been gathered in Mexico – a country where civic and ethical education forms a core part of the secondary school curriculum. Values, citizenship, moral education and other cognate areas have long-established educational traditions in Mexico, often linked to wider political agendas for peace, nation building and social cohesion.

Public schools are co-educational and serve the lower socio-economic classes of the population. By law, public schools must provide a secular ethical education, which promotes the principles enshrined in the Mexican constitution (Patiño-González 2009). This commitment to the state and the ethical principles mandated in citizenship is perhaps best symbolised in schools by the flag-honouring ceremony, conducted weekly – principles that are reinforced in weekly civics and ethics lessons. This mandatory curriculum includes topics such as education for democratic values, social and personal development, moral values and human rights.

Private schools, by contrast, are typically populated by students whose families can afford to have them attend, and this implies at least some social advantage. In the burgeoning Mexican private schooling sector – of which many have a religious character – developing the moral person is a marketable commodity. For example, in the city where these case studies were conducted, the Society of Mary (Marianos) advertised schools as giving formation in values (‘formación en valores’), as well as better opportunities for music, sport, dance, information technology and English. Private schools in
Latin America achieve higher academic results than public schools, although the attainment gap favouring private schools is smaller in Mexico than in some other Latin American countries, according to 2012 PISA (Program for International Student Assessment) scores (Geovanny, Gregorio, and Domingo 2017). These same authors conclude that academic attainment gaps are due to differences in individual, family and school characteristics with individual differences being most important.

In wider Mexican society, despite the secular constitution and its ramifications for education, there is a strong influence of the Catholic Church – with growing minority movements of evangelical protestant Christianity. Polls show over 90% of Mexicans were raised as catholic, with 81% retaining that identity (Pew Research Centre 2014). While the most recent census data suggest that almost 83% of the population in Mexico identify as Catholic (INEGI 2010). As with other countries in Latin America, the Catholic Church therefore has a considerable explicit and implicit influence on education (Klaiber 2009).

The potential impact of religion on adolescents’ lives is multifaceted – affecting issues as diverse as educational satisfaction and attainment (Mooney 2010). Research conducted elsewhere in Latin America suggests that religious commitment has a significant impact on adolescent behaviour, promoting pro-sociality and protecting against risk behaviours (author removed for anonymity). Exploring these factors among adolescents is important because understanding the impact of local moral norms has been a neglected issue in the field of character education internationally (Alexander 2016).

Comparing moral reasoning among students in two purposively sampled Mexican secondary schools provides suitable contexts to investigate the relationships between individual and social difference, schooling and adolescent moral character. Mexico, like many countries in Latin America, remains a starkly unequal society, particularly in terms of education and gender – despite having shown some signs of improvement in recent years (Esposito and Villaseñor 2017; Gasparini, Cruces, and Tornarolli 2011), and one manifestation of inequality is in the existence of public and private schools. Yet at the same time, in Mexico, values, moral and character education have long-established educational traditions. An embedded case study design presents a novel way to research relationships between moral character and social differences because it allows for a context-specific investigation into how various aspects of students’ lives, including their schooling, may be related to their ability to make moral judgements. We explore areas of specific interest regarding AD-ICM results: first, for the larger public-school sample we investigate connections between students’ AD-ICM results, family and school, in addition to differences between boys and girls, and second we investigate differences between the girls’ scores across the two schools.

**Adolescents’ moral reasoning**

Developed in the neo-Kohlbergian tradition, the intermediate concept approach (responding to criticisms of Laurence Kohlberg’s theory) is concerned with cognitive features of virtues as ‘intermediate concepts.’ These are believed to exist between two other defined levels of moral understanding as concepts that are specific to daily life and contexts (Rest et al. 1999; Thoma, Derryberry, and Crowson 2013). They are located between highly contextual norms which prescribe action in very specific circumstances (moral codes and regulations for example) and more comprehensive default interpretive
systems (bedrock schema) such as personal interest, maintaining norms or post-conventional schemas (cf. Rest et al. 1999). The foundational assumption of the ICM measure, as part of this tradition, is that patterns of ratings and rankings in participants’ responses to dilemmas reveal reliable information about their ability to interpret and apply virtue (or intermediate concepts) in the real world. While this omits assessment of moral emotion or behaviour, two other commonly assumed components of moral character, it provides a means of exploring individual and contextual factors in adolescent character development about one key aspect – the ability to make appropriate moral judgements (Berkowitz 2011). Thus, AD-ICM is able to measure adolescents’ moral reasoning as related to the application of virtues to real-life situations relevant to the context of adolescence and school life. Although the emphasis of the measure is on context as a stage of life (adolescence) with the dilemmas representing likely or feasible situations an adolescent might face relating to school, context is relevant too in terms of culture. Little is known about the application of the virtues across cultural contexts, although evidence exists for at least some cross-cultural similarities of the virtues (McGrath 2014; Niemiec 2013; Thoma et al. 2019). Similar to virtues, intermediate concepts assessed by AD-ICM (honesty, courage, self-discipline) may have some cross-cultural similarities. A recent cross-cultural comparison of adolescent responses to AD-ICM (Thoma et al. 2019) found both similarities and differences for the application of virtues, depending on (among other factors) whether a country is collectivist or individualistic (Hofstede 2001). Patterns of response to the dilemmas across cultures were similar, even though measure development took place in US with a local expert panel.

While AD-ICM has proved to be a useful measure for viewing the application of virtues across cultural contexts, previous application of AD-ICM within a specific culture suggests that certain individual and contextual factors have significant correlations with adolescents’ moral judgement ability. For example, in a large UK study, being female and practising a religion were positively correlated with high scores (Arthur et al. 2015; Walker et al. 2017). Schools’ average scores were also found to differ on the AD-ICM suggesting that some educational and contextual settings may be more or less successful in encouraging students’ moral development. To explore these findings further researchers first focused on the kind of school (e.g. religious, public, private, etc.) as a proxy for different educational climates. These analyses suggested that high performing scores could be found across the common categories that define UK schools, although certain socio-economic factors were associated with lower AD-ICM results. To help clarify the characteristics of high performing schools, researchers explored more nuanced school-related characteristics such as students having parents that did/did not go to university, going to a school with a higher percentage of free school meals’ eligibility or in an area with high unemployment rates, all of which showed small but significant negative correlations with AD-ICM results. Researchers in the same study found correlations with high scores and practices of character education (such as having at least one teacher who is especially passionate, knowledgeable and skilled in developing students morally). However, they were unable to explain fully why some schools achieved higher average scores than others. The purpose of the present study is therefore to explore the relationships between individual and contextual factors and moral reasoning in two well-defined and distinct educational contexts.
Methods

Drawing on previous research using AD-ICM, we set out to explore if there were differences in moral reasoning between adolescents from very different social backgrounds who attended different kinds of schools. We were particularly interested in gender and religion as previous studies have shown girls, and adolescents who practice a religion have significantly higher scores on the AD-ICM (Walker et al. 2017). We chose an embedded case study design that would allow for further insights into character development processes taking place in schools. Our guiding research questions can be stated as follows:

- What individual differences are related to moral reasoning ability?
- What social differences are related to moral reasoning ability?
- What differences in educational provision may relate to differences in moral reasoning?

In order to examine these questions, two case study sites were selected, using a purposeful diverse-case sampling strategy: one comparatively well-performing public secondary school with a dedicated values education program, and one girls’ private school with a distinct virtue-based character education program. These two sites represent diverse cases whereby the variables of socio-economic status of family, school resources, moral education program, religious ethos and single-sex schooling can be examined in context. Ethical scrutiny followed the procedures of the University of Navarra under the principles of the American Educational Research Association guidelines (AERA 2011). Participants’ provided voluntary informed consent and parents were asked if they wanted to exclude their child from the study. Data generation was voluntary, confidential, and anonymous, and students could withdraw from the research at any time without needing to state a reason.

Participants

With the aim of maximum participation, we invited all students from both schools to consent and take part in the research. From Colegio Privada, 298 girls took part in the study. This represents 58% of the school population at the time. Ninety-five percent of participating students said that they practised a religion. From Secundaria Pública, 988 students took part. This is 62% of the school population at the time. Of participating students, 62% were female and 68% said they practised a religion.

Measure

Students completed an electronic version of the Adolescent Intermediate Concept Measure (AD-ICM) in March 2017. A reduced version of this survey was used because the additional time required to complete all seven dilemmas was too disruptive for school routines (2017). The shortened version retains the structure of the original AD-ICM measure (Walker et al. 2017), but comprises only three moral dilemmas. Each dilemma emphasises a particular virtue or intermediate concept: honesty (should one report...
a cheating incident involving peers in a school context?), self-discipline (whether to attend a final-year trip or prepare for a maths exam?) and moral courage (should one stand up to a gymnastics coach to uphold personal values?). Participants read each dilemma and rated approximately ten action choices and ten reasons/justifications on a scale from 1 (‘I strongly believe that this is a GOOD choice/reason’) to 5 (‘I strongly believe that this is a BAD choice/reason’). Following the rating task, participants select and rank (first, second and third) best and worst (and second worst) options for actions and reasons as the basis for scoring. The survey was translated into Spanish and was piloted with a small group of students in one school to ensure its comprehensibility.

The intermediate-concept approach is underpinned by the theoretical assumption that when related to specific contexts (e.g. professional groups or in this case a stage of life: adolescence), we can generally agree on options that are appropriate and inappropriate, but we will struggle to agree on precisely what a person ought to do in certain real or hypothetical situations if they are to exercise good moral reasoning (Rest et al. 1999). From the perspective of the intermediate concept approach, this involves judgements about what should be done and reasons for acting, including the identification of poor options. A full account of this approach (including an overview of the expert panel process and the extensive development and testing process for ICM) is available (Thoma, Derryberry, and Crowson 2013). AD-ICM results, reported below, are easily understood as percentage scores representing to what extent participant choices match prior expert panel definitions of each item as either acceptable, unacceptable or neutral. For example, participants selecting as good three ‘acceptable’ action choices for the protagonist will score maximum points for that part of the survey. A participant with an overall score (Total ICM) of .80 is making action and justification choices in line with the expert panel at a level of 80%. Importantly, there are multiple ways to score highly meaning that moral reasoning is assessed rather than a participant’s choosing of a particular course of action.

In addition to the three dilemmas, the questionnaire also recorded participants’ self-reports of the following variables: gender; age; which family members they lived with; if they practiced a religion; if they played sport; and, if they thought their school grades were better or worse than their peers. In addition to AD-ICM scores, information about each school’s moral and character education provisions was recorded in fieldwork notes.

As a relatively new measure, knowledge about how far AD-ICM can be sensitive to different cultures is only beginning to emerge. In each of four countries outside of USA where the measure has been used – including this study – the process of adjusting the measure was straightforward suggesting cultural transferability for the measure. That is, (a) the translation process was not difficult because concepts were easily recognisable to translators; (b) pilot testing showed the dilemmas were pertinent and lifelike for participating adolescents – the stories made sense to them, and (c) the behaviour of the measure as used in data collection across cultures supports its viability in that there were not large numbers of violations or a lot of missing data, etc. Additionally, sensitivity to cultural variation for AD-ICM is assessable by carefully attending patterns of choice made by respondents from different cultures or contexts (e.g. a public and private school in Mexico). So far, this has been done to the extent that choices attract better or worse results in comparison to expert panel choices. For example, the measure favours choices involving varying degrees of action towards good outcomes over inaction/conforming to the group (i.e. taking a stand against cheating versus conforming to peers and being
silent). Understanding these culturally different AD-ICM response patterns is in its infancy, and this article represents a small step in this regard. However, there remains a real need to incorporate better understanding of cultural difference in interpreting measures such as this one. For example, participant dilemma choices not to take action and to conform to peers probably implicate broader cultural and social forces in addition to ethical theory. For example, adolescents routinely experiencing negative consequences for making any personal stand seem unlikely to do so on the measure, and so future effort to account for contextual influences such as this is important for interpreting AD-ICM. Equally, we do not yet know how severe material deprivation might impact responses to measures such as this. Overall, further exploration of social and cultural influences on responses to the measure (e.g. in individualistic versus collectivist countries) are needed to fully understand the extent of its cultural sensitivity.

Findings

Overview of the two schools

Located approximately 11 km apart, the schools represent considerably different social, ideological and educational contexts. Colegio Privada is a private girls’ school with a Catholic ethos with an elementary, junior high and high school on the same campus. At the time of the fieldwork, there were 509 girls on roll, 134 in grades 6–7 (equivalent to grades 1–3 in Secundaria General). Girls attended the school in one shift starting in the morning and received instruction in English and Spanish. Secundaria Pública is a co-educational public junior high school catering for grades 1–3 of Secundaria (ages 12–15). It is situated in a lower socio-economic status neighbourhood and operates in two shifts, one in the morning (7 until 13.20) and one in the afternoon (13.30 until 19.40). At the time of the fieldwork, the morning shift had 857 students on roll and 741 in the afternoon. While all of the differences observed between the schools are not possible to detail nor explore in this article, Table 1 (below) summarises some key observed differences in the characteristics of the schools.

Moral education programs in the two schools

Both schools have dedicated moral education programs in addition to teaching mandatory discrete Formación Cívica y Ética (Citizenship and Ethics) classes weekly as required by Mexican law. Secundaria Pública identified eight values of focus for students and teachers (referred to as ‘valores’ values): responsibility, discipline, cleanliness, solidarity, perseverance, respect, commitment and punctuality. These values are very much aligned with conformity, even control. Purpose-made banners displayed these values prominently throughout the school, including at the school entrance. Over the course of an academic year, one month was given over to each value. During this time, teachers in each subject area linked aspects of their classes to the value of the month. At the same time, various readings are selected for the whole school that illustrate the same value – excerpts of which are read by a student daily through a public address system. At the time of the research, the school was listening to a Spanish translation of Be a Happy Teenager (Andrews 2002). Students also take part in an awareness campaign creating their own posters and slogans.
Table 1. Key observed differences in the characteristics of the schools.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Secundaria Pública</th>
<th>Colegio Privada</th>
</tr>
</thead>
<tbody>
<tr>
<td>Religion</td>
<td>Secular (no organised religious observances); no religious instruction</td>
<td>Catholic (daily voluntary Mass); discrete religious education classes</td>
</tr>
<tr>
<td>Approximate median class size</td>
<td>48</td>
<td>21</td>
</tr>
<tr>
<td>Overall student population</td>
<td>Total roll 1598</td>
<td>Total roll 509</td>
</tr>
<tr>
<td>Student and teacher gender</td>
<td>Co-Educational (male and female teachers)</td>
<td>Females only (teachers and students)</td>
</tr>
<tr>
<td>Student truancy</td>
<td>Students observed out of lessons</td>
<td>No truancy observed</td>
</tr>
<tr>
<td>Language of instruction</td>
<td>Spanish</td>
<td>Spanish and English</td>
</tr>
<tr>
<td>Security</td>
<td>Gate keeper</td>
<td>Private security guards and armed police escort</td>
</tr>
<tr>
<td>Geographical location</td>
<td>In urban commercial and residential district in the city on low ground</td>
<td>In suburban residential district above the city on high ground</td>
</tr>
<tr>
<td>Funding</td>
<td>Public (free)</td>
<td>Private (fees undisclosed)</td>
</tr>
<tr>
<td>Extra-curricular activities (during the school day or after school)</td>
<td>Unstructured/conducted by volunteer teachers, including literature and basketball</td>
<td>Wide range of structured activities, including dance, athletics, cooking</td>
</tr>
<tr>
<td>Student behaviour</td>
<td>Some students off task in lessons/bullying (reported)</td>
<td>Some students off task in lessons/bullying (reported)</td>
</tr>
</tbody>
</table>
Colegio Privada had a more personalised and systematic approach to character development aimed at cultivating a personalised harmony of virtues for each individual student, as opposed to espousing desired values. While identified values and virtues were often the same across schools (e.g. solidaridad (solidarity) and respeto (respect) feature in school directives), Colegio Privada placed these into an overarching framework of virtues, with bodily, affective, intellectual and motivational dimensions. Classes in Formación Cívica y Ética in Colegio Privada thus followed a curriculum giving greater prominence to virtues and to individual cultivation of character or autonomy compared to Secundaria Pública where there is an emphasis on values aligned with conformity. In addition to this, students were individually mentored within the same framework, which aimed to cultivate each individual’s holistic development across behavioural, affective and intellectual domains. This framework was a prominent part of each teacher’s approach to their own subject, and part of the extensive extra-curricular activities which were coordinated centrally by a curriculum advisor. Consequently, teachers and students in Colegio Privada reported being proud of the school and its role in character development. On the other hand, while some teachers in Secundaria Pública were intensely involved in establishing a whole school values program (which demonstrated their dedication and belief in the intrinsic worth of their work), students and teachers were more ready to identify problems with the school, its resources and other students and staff.

Results of the intermediate concept measure

Secundaria Pública
As shown in Table 2, students at Secundaria Pública performed on average as well as students taking the survey in other studies, achieving almost a 50% match with expert panel choices (.48), with results evenly distributed across percentiles (25th = 0.30; 50th = 0.52; 75th = 0.70). Also, similar to other studies (cf. Walker et al. 2017), results for the four subscales that make up the main score (total ICM) vary. Inspection of the means and associated standard errors for the subscales indicates that participants found it easier to make best selections (both action (M = 0.55) and justification (M = 0.49) choices) rather than worst selections (action (M = 0.41) and justification (M = 0.37) choices), suggesting students found it more difficult to identify poor options for what the protagonist should do and why. As for identification of the protagonist’s best options, scores were higher for best action than for best justification choices (0.55 versus 0.49). These within-subject differences are indicated by a significant main effect using the Greenhouse–Geisser correction for the absence of sphericity \[F(2.66,6.57) = 71.58; p < 0.001; \eta^2 = 0.076\].

<table>
<thead>
<tr>
<th>Table 2. Total ICM scores including gender.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variable</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Standard deviations are included in parentheses.
As in other studies (Walker et al. 2017), girls outperformed boys overall (M = 52 versus M = 43) \([F(1,1.50) = 20.94, p < .001, \eta^2_p = .023]\) and for subscales, where a between subject main effect was found \([F(1,4.88) = 16.14, p < .001, \eta^2_p = .018]\). This main effect was conditioned by a gender-by-subscale interaction effect \([F(2,7,0.316) = 3.45, p < .001, \eta^2_p = .018]\) which, though small, suggests within-subject differences between subscales. Inspection of the means and standard deviations shows larger differences between subscales for girls, especially between their best and worst selections, such that girls seem disproportionately adept selecting good than bad selections.

A gender by age interaction effect was also found for total AD-ICM scores \([F(2,0.21) = 3.01, p < .05, \eta^2_p = .007]\). For these analyses, students were sorted into three age groups: 12, 13 and 14 &15 (see Table 3). Ages 14 and 15 were combined to achieve a sufficiently large group. Inspection of means and standard deviations (Table 2) suggests total AD-ICM scores increased for girls with age, but apart from an increase between ages 12 and 13, scores for boys declined with age thereafter. Gender differences are most marked for students aged 14 and 15, with girls achieving significantly higher total AD-ICM scores than the boys (M = .58 versus M = .43). Relating to the subscales, although no within-subject differences were found, unsurprisingly there was a between-subject gender by age interaction effect \([F(2,0.98) = 3.29, p < .05, \eta^2_p = .008]\) showing a similar pattern of results, except that boys aged 13 performed much better than girls of the same age for worst action choices (M=.47 versus M = .43), but dropped well below girls in the age group 14 and 15 (M = .37 versus M = .49).

How did students in Secundaria Pública respond to each of the dilemmas? As shown in Table 4, students scored highest for the dilemma assessing self-discipline (M = .57), followed by honesty (M = .46) then courage (M = .41). This is shown by a significant within-subject

### Table 3. ICM scores by age, gender and ICM subscales.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age</th>
<th>Sample</th>
<th>Total ICM</th>
<th>Action choices</th>
<th>Justification choices</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Best</td>
<td>Worst</td>
</tr>
<tr>
<td>Male</td>
<td>All</td>
<td>399</td>
<td>.44(.27)</td>
<td>.51(.32)</td>
<td>.40(.44)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>118</td>
<td>.41(.26)</td>
<td>.48(.29)</td>
<td>.36(.44)</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>151</td>
<td>.47(.28)</td>
<td>.52(.34)</td>
<td>.46(.39)</td>
</tr>
<tr>
<td></td>
<td>14/15</td>
<td>130</td>
<td>.43(.29)</td>
<td>.51(.31)</td>
<td>.37(.48)</td>
</tr>
<tr>
<td>Female</td>
<td>All</td>
<td>466</td>
<td>.52(.26)</td>
<td>.59(.30)</td>
<td>.43(.43)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>164</td>
<td>.48(.28)</td>
<td>.56(.34)</td>
<td>.38(.44)</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>192</td>
<td>.51(.25)</td>
<td>.58(.29)</td>
<td>.43(.44)</td>
</tr>
<tr>
<td></td>
<td>14/15</td>
<td>110</td>
<td>.58(.23)</td>
<td>.65(.26)</td>
<td>.49(.38)</td>
</tr>
</tbody>
</table>

Standard deviations are included in parentheses.

### Table 4. Total ICM scores by dilemmas, age and gender.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age</th>
<th>Self-Discipline</th>
<th>Honesty</th>
<th>Courage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>all</td>
<td>.57(.41)</td>
<td>.46(.37)</td>
<td>.41(.33)</td>
</tr>
<tr>
<td>Male</td>
<td>all</td>
<td>.52(.44)</td>
<td>.40(.39)</td>
<td>.38(.31)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>.47(.44)</td>
<td>.39(.37)</td>
<td>.36(.29)</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>.55(.44)</td>
<td>.45(.37)</td>
<td>.39(.34)</td>
</tr>
<tr>
<td></td>
<td>14/15</td>
<td>.52(.43)</td>
<td>.37(.42)</td>
<td>.38(.31)</td>
</tr>
<tr>
<td>Female</td>
<td>all</td>
<td>.61(.38)</td>
<td>.51(.35)</td>
<td>.44(.33)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>.55(.41)</td>
<td>.50(.36)</td>
<td>.41(.34)</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>.62(.38)</td>
<td>.49(.36)</td>
<td>.41(.34)</td>
</tr>
<tr>
<td></td>
<td>14/15</td>
<td>.68(.37)</td>
<td>.54(.31)</td>
<td>.51(.30)</td>
</tr>
</tbody>
</table>

Standard deviations are included in parentheses.
main effect, correcting for the absence of sphericity with the Greenhouse-Geisser test
\[F(1.93,5.92) = 59.26, p < .001, \eta_p^2 = .063\]. Once both age and gender were taken into account, there was a significant between-subject interaction effect \[F(2,0.64) = 3.01, p < .05, \eta_p^2 = .007\]. This repeats the pattern of response to the dilemmas described above (self-discipline, honesty and moral courage), except that girls’ scores increased with age for each dilemma, whereas boys’ scores increase only between the age groups of 12 and 13. Gender differences by dilemma, favouring girls, are again largest for the group aged 14 and 15, especially for the dilemma about honesty (\(M = .54\) versus \(.37\)) – a dilemma that in some respects pitches honesty against loyalty to one’s peers.

**Other factors.** Factors with a statistically significant relation to AD-ICM results for this school were: family, religion, perception of grades compared to peers and taking part in sport. These results are described below, beginning with a description of the students’ family arrangements (see Table 5).

**Family.** Ninety-five percent of participating students from Secundaria Pública lived with at least one of their parents, whereas only 68% lived with both parents. Eighteen percent of the students lived with their grandmothers and 76% of the students lived with one or more sibling. Some students had others living with them such as an uncle or niece (9%). Students that lived in a household with their grandmothers achieved higher ICM results than students that did not (\(M = .54\) versus \(M = .46\)) \[F(1,0.780) = 10.750, p < .001, \eta_p^2 = .012\]. Similarly, students living in a household with a sibling(s) also achieved higher results than those that did not (\(M = .49\) versus \(M = .44\)) \[F(1,0.95) = 5.411, p < .05, \eta_p^2 = .006\]. Apart from 18 students that lived only with their grandmothers, students living with siblings or grandmothers (or both) also had at least one parent in the household. Students scored better than those in other living arrangements if they lived in the following company: at least one parent, a sibling or more and their grandmothers (\(M = .55\) versus \(M = .47\)) \[F(1,0.808) = 11.132, p < .001, \eta_p^2 = .013\]. The presence in the household of individuals other than these such as an uncle or niece, for example, did not significantly influence ICM results. These main effects were not conditioned by gender.

| Table 5. ICM scores by living arrangements, religion, sport and grades. |
|-------------------|-----------------------|---------|----------|
| Variable          | Options                            | Sample  | AD-ICM   |
| Family            | Live in household including grandmother | 167     | .54(.24) |
|                   | Live in household without grandmother | 708     | .46(.27) |
|                   | Live in household with brother and or sister | 663     | .49(.27) |
|                   | Live in household without brother or sister | 212     | .44(.27) |
|                   | Live with at least 1 parent and 1 sibling, plus grandmother | 126     | .55(.24) |
|                   | Other living arrangements           | 749     | .47(.27) |
| Practise religion | Yes                                | 598     | .49(.27) |
|                   | No/don’t know/rather not say        | 276     | .45(.28) |
| Sport             | Yes                                | 577     | .47(.27) |
|                   | No                                 | 298     | .50(.27) |
| Grades            | Above/mostly above peers           | 318     | .51(.27) |
|                   | About same/below peers             | 548     | .46(.27) |

Standard deviations are included in parentheses.
Religion. Students were asked if they practised religion or not. Those claiming to practise religion scored higher on AD-ICM than those who said that they did not practise religion, did not know or preferred not to say (M = .49 versus M = .45) [F(1,0.35) = 4.719 p < .05, \( \eta_p^2 = .005 \)]. These main effects were not conditioned by gender.

Grades. Students differed on their AD-ICM scores regarding the perception of their own school grades compared to their peers [F(1,0.61) = 8.29; p < .005, \( \eta_p^2 = .006 \)]. Post-hoc comparisons indicate that highest AD-ICM scores were associated with those claiming to have better or mostly better grades (M = 0.51), compared to those claiming to have grades ‘about the same’ or ‘below’ their peers (M = 0.46). These main effects were not conditioned by gender (See Table 8).

Sport. While we did not ask which sports students were doing, most of the students in the study claimed to do sport and as shown in Table 4, these students performed less well for AD-ICM than those that said they did not do sport (M = .47 versus M = 50) [F(1,0.29) = 4.00 p < .05, \( \eta_p^2 = .005 \)]. In the UK study, though not statistically significant, there was a trend towards similar results (Author removed for anonymity).

Colegio Privada
Next, schools were compared by excluding boys from analyses to explore differences between girls’ responses to AD-ICM. School samples were quite similar in terms of the girls’ average age (Colegio Privada: 14.6 (2.33); Secundaria Pública: 12.97 (.93)) and for the number of girls living with at least one parent (99% in Colegio Privada and 95% in

<table>
<thead>
<tr>
<th>Variable</th>
<th>Options</th>
<th>Sample</th>
<th>AD-ICM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grades</td>
<td>Above/Mostly above peers</td>
<td>139</td>
<td>.64(.22)</td>
</tr>
<tr>
<td></td>
<td>About same/Below peers</td>
<td>130</td>
<td>.55(.25)</td>
</tr>
</tbody>
</table>

Standard deviations are included in parentheses.
Secundaria Pública). But, of course, schools differed by class and affluence and this shows across some variables. For example, in Colegio Privada, 78% of students had both parents go to university compared to 21% in Secundaria Pública. In family households, far more female students in Secundaria Pública (20%) lived with their grandmothers than in Colegio Privada (5%). Interestingly, 52% of girls in Colegio Privada self-identified as having grades above or mostly above others, compared to 37% in Secundaria Pública, perhaps reflecting higher levels of confidence. Schools were also dissimilar regarding the percentage of girls claiming to practice religion. Indeed, 95% of students in Colegio Privada claimed to practice religion compared to only 69% in Secundaria Pública. Given, the connection between better AD-ICM results and practising religion found in other studies, this variable was tested in order to assess whether results described below were altered by this imbalance. This was done by excluding students in Secundaria Pública who said they did not practice religion. Results reported below were not disrupted by this exclusion.

Inspection of Table 6 shows that girls in Colegio Privada achieved significantly higher total AD-ICM results than girls in Secundaria Pública \[F(1,0.82) = 12.36, p < .001, \eta_p^2 = .019\]. Breaking this down for the subscales, there were statistically significant differences between, but not within, subjects \[F(1,3.00) = 10.79, p < .001, \eta_p^2 = .017\]. Inspection of the means and associated standard deviations in Table 5 shows that girls in Colegio Privada scored higher than girls in Secundaria Pública across all of the subscales, but especially for justification choices with differences of 12 and 9 percentage points for best and worst choices, respectively.

In comparing next how girls in the different schools responded to specific dilemmas, significant between-subject differences were found \[F(1,2.46) = 12.36, p < .001, \eta_p^2 = .019\]. Inspection of means and standard deviations in Table 7 shows girls in Colegio Privada scored much higher for dilemmas involving self-discipline and moral courage, but were almost equal to girls in Secundaria Pública for responses to the honesty dilemma. Significant within-subject differences were also found for the dilemmas, again correcting for the absence of sphericity using the Greenhouse–Geisser test \[F(1.99,0.75) = 8.52, p < .001, \eta_p^2 = .013\]. However, according to post hoc comparisons, these within-subject differences were statistically significant only between the dilemmas for self-discipline and honesty and between self-discipline and moral courage, but not between honesty and courage.

Testing if total ICM scores were related to other variables such as family living arrangements, religion, etc., for girls at Colegio Privada was not possible, but sample sizes were sufficient to test how perception of school grades related to scores, and those students claiming to have grades above or mostly above their peers scored higher for Total AD-ICM (M = .64 (.21)n = 88) than those students who perceived their grades as the same or below peers (M = .55 (.28)n = 84) \[F(1,0.37) = 10.11 p < .05, \eta_p^2 = .035\].

**Discussion**

Regarding individual differences in moral reasoning, gender was a significant factor of variance for responses to the dilemmas. For Secundaria Pública students, girls outperformed boys in all ways. This was not surprising because females commonly achieve higher results on moral and character tests (e.g. Malti and Buchmann 2010; Nunner-Winkler 2007).
Furthermore, while girls consistently achieved higher scores than boys, when age was taken into account gender difference for AD-ICM was especially large in the age group 14/15. It is noteworthy, too, that boys aged 13 were uncharacteristically adept at identifying worst action choices, perhaps indicating they were coming into awareness of what older boys were doing (they scored lowest of all). Gender differences were smallest for the dilemma about moral courage, and once age was taken into account, the largest gender difference occurred for the dilemma about honesty in the age group 14/15. These differences at ages 14/15 do not account for girls doing better overall, since if we remove this age group girls still do significantly better. Given the pull in the honesty dilemma between loyalty to friends/peers and doing the right thing by making a stand against cheating, further research into processes of social interaction among peer groups in this age group may provide some explanation here. As it is, there are a number of tentative possibilities for depressed honesty scores. Perhaps in the interim adolescent period between dependence-on and independence-from parents (Steinberg and Silverberg 1986), peer influence is stronger (Steinberg and Monahan 2007) involving both increased salience of the peer group and an increased need to belong (Brown, Clasen, and Eicher 1986). As for gender differences for this dilemma, possibly peer/friendship groups among boys at this age place greater significance on solidarity to peers/friends than girls do, since girls seem less influenced than boys by peer groups generally (Berndt 1979; Steinberg and Silverberg 1986). Males also seem less likely to conform to socially desirable actions and justifications (Bernardi 2006; Bernardi and Guptill 2008) and may even be more likely to cheat themselves (Borkowski and Ugras 1998) with implications for their tolerance of cheating in the dilemma. These tentative explanations indicate that individual differences in moral reasoning may have at least some complex socially determined aetiologies. Cultural context at the country level is likely to be relevant here too. When responses from the students in Mexico are compared to responses from other countries with individualistic cultures, we can see that Mexican students more readily choose loyalty to peers/friends in the honesty dilemma. This suggests that low honesty scores may be a partial expression of belonging to a collectivist culture where there is greater emphasis on group needs (Hofstede 2001; removed for anonymity). As for courage, boys are closer to girls when responding to this dilemma which involves girls and gymnastics, requiring courage to stand against authority (the coach). Perhaps a rather undeveloped non-conformist tendency (in the sense that it is not admirable or morally oriented) among the boys is working both for and against them in the dilemmas? Other kinds of response are triggered by the self-discipline dilemma, which may be prompting responses from boys and girls in rather more stereotypical directions, and reflecting real-time descriptions of the students provided by their teachers: on average, boys show less self-discipline than girls.

Some discernible social differences are also related to variances in moral reasoning. Students in Secundaria Pública living in a household with their grandmother and or sibling(s), together with at least one parent achieved higher AD-ICM scores, as did students practising religion and claiming that their school results were above or mostly above peers. Conversely, doing sport was associated with lower AD-ICM results. These findings could be explained in recourse to several tentative theories meriting further investigation. Large families may provide greater opportunities for social interaction from an early age, impacting on the development of moral reasoning, for example.
Practising a religion may give the opportunity for moral reflection, and present adolescents with role models (Sanderse 2013) for exemplary moral behaviour, as well as demanding particular codes of conduct, such as telling the truth and treating others with respect. Research conducted elsewhere in Latin America suggests that religious commitment has a significant impact on adolescent behaviour, promoting pro-sociality and protecting against risk behaviours (Moulin-Stożek et al. 2018b). Likewise, the inverse relationship between practising sport and moral reasoning is interesting. Could it be that practising sport consumes free time that could otherwise be spent in more reflective activities such as reading good literature that provide greater opportunities to develop moral reasoning ability? Alternatively, according to local observations students practising sport in Secundaria Pública may be more likely to spend time on the streets or truanting from lessons in order to play sport, indicating different priorities and social practices in comparison to their peers.

The findings also suggest that differences in the educational provision in the schools may account for differences in moral reasoning. It is especially notable that the private school aims to cultivate self-leadership and integrated character while the values base approach of the public school is concerned with conformity, possibly even control. While it is difficult to differentiate between the effects of school and of social class, it seems likely that at least some of the differences observed are relatable to the different school approaches to moral and character education. Quite clearly, girls in Colegio Privada achieved higher Ad-ICM scores overall than girls in Secundaria Pública. Better performance was particularly noticeable for making justification choices and for the dilemmas assessing self-discipline and courage over honesty. Self-perception of higher educational attainment compared to peers was related to AD-ICM scores for Colegio Privada students as it was for students in Secundaria Pública, but interestingly, far more Colegio Privada girls identified themselves as having grades above or mostly above their peers. It is telling that beyond achieving different score levels, the pattern of responses to the three dilemmas was more similar between boys and girls from Secundaria Pública than between girls from Colegio Privada and girls from Secundaria Pública. For example, Colegio Privada girls scored very well for self-discipline and courage, compared to honesty, whereas boys and girls in Secundaria Pública achieved higher scores for self-discipline, then for honesty, closely followed by moral courage. It is likely that the conception and experience of the virtues of self-discipline, honesty and courage in Colegio Privada are different to how girls and boys in Secundaria Pública conceive and apply them? Differences in moral education provision do seem to correlate to moral reasoning but of course, this is confounded by the very different social backgrounds of the girls. It is revealing that girls in Colegio Privada were more likely to say they are better than their peers at school work than girls from Secundaria Pública. Higher courage and self-discipline scores also suggest these girls may be more confident and able to voice their own needs. Given that role modelling is a fundamental aspects of character education theory, perhaps girls in a single-sex environment where independent, female role models are readily available, are more likely to stand up for themselves and say what they believe, or more stereotypically this could also be down to these tendencies being more common among students in higher social classes?
choices, since this is generally considered a higher level moral agency. Further research is needed into these social processes.

**Limitations**

This exploratory case study has some limitations. Firstly, as a cross-sectional study between two schools, differences by school type, though statistically significant, remain difficult to pinpoint. However, finding greater similarity between boys and girls in one school versus the girls from two schools is suggestive of school and class-based differences for the application of virtue to the dilemmas. Secondly, given a smaller sample size for the girls’ school, it was not possible to run a complete suite of analyses between age groups according to gender, nor to make comparisons between boys in different school types. Further investigation is warranted because we cannot be sure if the single-sex environment of Colegio Privada contributed to the difference in performance between cohorts of girls, or whether this could be due to other factors, such as the socio-economic status, or education of the parents. We also, of course, cannot surmise if the same effects would also be demonstrable between cohorts of boys between mixed and single-sex schools.

**Conclusions**

Proponents of character education claim the cultivation of virtues during schooling helps students, schools and society to flourish. However, critics argue that character education programs are insensitive to cultural and social diversity, and implicitly justify inequality by accrediting it to the presence or absence of personal qualities which are in principle unmeasurable. The present study, though small scale, provides some contextualised evidence relevant to this debate. There were significant differences in performance in moral reasoning between girls attending the two different schools. Although from the present study it is not possible to establish if it is schooling or social background that have the most salient impact on the ability of adolescents to apply virtues to the specific dilemmas, it looks as if social inequality may well bear on character outcomes. Most obvious here is the difference in performance between the cohorts of girls in public and private sector schools. In addition to schooling, family living arrangements and religion also appear to be significantly related to moral reasoning, suggesting that critics of character education are justified in arguing that social inequality and differences in cultural norms may confound character education efforts, giving advantage to particular demographic groups. Even so, in terms of ‘moral cultivation of youth’, we believe it is heartening that engaging in events and practices related to virtue cultivation (such as participating in religious activities and being part of a family incorporating multiple generations) are associated with higher levels of moral reasoning. However, differences in moral reasoning for girls by social background and school do stand out as disadvantaging girls from Secundaria Pública who are drawn to choices of conformity over those of empowerment in responding to dilemmas. Rather than casting a shadow on character education, we view this as evidence of need for variants of character and moral education developing critical reasoning, especially for appropriately challenging norms and social convention, a key feature of highest levels of moral reasoning according to neo-
Kohlbergians (e.g. post-conventional reasoning). The diverse-case strategy used in this research does, however, suggest several variables for testing in future studies based on key differences between these schools, such as class size, presence or absence of individual mentoring, resources available for the character education curriculum, and perhaps most controversial of all, the possible effect of the single-sex environment, not to mention approaches to cultivating character that work towards empowerment versus conformity.

Aside from this debate, one important conclusion of the study – and one respect where scholars may be exaggerated in their criticisms of character education – is the potential to reliably measure and evaluate virtues – in this case the application of the virtues to specific contexts as a form of moral judgement, so important for good character. The AD-ICM is a measure of moral judgement capable of identifying variances in moral reasoning among different cohorts and across cultural and linguistic divides with a satisfactory degree of reliability (Thoma et al. 2019).

Acknowledgments

This research was supported by a grant from the Templeton World Charity Foundation, 'Researching and promoting character education in Latin American secondary schools' (0157), held by the University of Navarra 2015–2018.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Templeton World Charity Foundation [0157].

ORCID

Daniel Moulin-Stožek http://orcid.org/0000-0002-9975-2530

References


