

Running head: Adolescents' application of the virtues

Thoma, S. Walker, D.I., Chen, Y.H., Frichand, A., Moulin-Stožek, D. & Kristjánsson, K

In press, Developmental Psychology, 2019

Assessing adolescent's application of the virtues across multiple cultural contexts

Abstract

Little is known about adolescent applications of the virtues such as honesty, responsibility and courage across different cultural contexts. Using the Adolescent Intermediate Concepts Measure we analyze samples of adolescents (ages 12-20 and N= 9112) from five contexts: The Republic of Macedonia, Mexico, Taiwan, The United Kingdom and The United States. Across samples, adolescents provide evidence of developmental growth in the ability to apply virtue concepts as assessed by responses to dilemma-based situations. Within these trends, participants found it easier to identify action choices that reflect the virtue concepts as compared to justifications for possible actions. Additionally, participants were better able to identify appropriate applications of the virtues as compared to inappropriate ones. Gender differences favoring females were noted across samples. Overall, similarities across settings were more striking than differences suggesting that there is value in viewing the virtues as a normative component of character development across the adolescent years.

Keywords: virtues, moral development, culture.

Assessing adolescents' application of the virtues across multiple cultural contexts

One may also observe in one's travels to distant countries the feelings of recognition and affiliation that link every human being to every other human being.

Aristotle, Nicomachean Ethics, 1155a21-2

Virtues have been used by philosophers and social scientists to explain a range of human characteristics including the foundations of human flourishing, the elements of character, and the targets of moral education (Seligman, 2011). To so-called virtue theorists the virtues are “...understood as settled states of character, concerned with praiseworthy conduct in significant and distinguishable spheres of human life” (Kristjánsson, 2015). Perspectives such as these assume at least some similarity across cultures and contexts for the identification and understanding of virtue. Whether this similarity is due to historical and cultural interactions between Western, Eastern and Asian cultures and ethical texts, the expansion of Western educational practices, or something fundamentally universal about what constitutes human flourishing, there is now growing support for viewing at least some of the virtues as having broad cross-cultural relevance and recognition (e.g. McGrath, 2014; Niemiec, 2013).

Although some virtues may be recognized and endorsed in a variety of cultural settings, it is less clear if the reflective application of the virtues to specific social problems also display common elements. It may be that individuals in various contexts have similar understanding of the virtues and yet have a setting-specific interpretation about how a virtue is best applied in real-life situations. To address this possibility we focus on the application of the virtues and assess whether moral dilemmas identified by adolescents in the United States (US) as relevant and

realistic, are evaluated in similar ways within four additional, and distinct, cultural settings: The United Kingdom (UK), The Republic of Macedonia (Macedonia), Mexico and Taiwan.

Although hardly an inclusive listing of cultural settings, these samples do provide information about virtue applications from cultures that differ in their social organization along a broad individualistic and collectivist dimension (e.g., Hofstede, 2001). In these broad-based conceptions of cultural patterns, individualistic societies prioritize the individual, personal goals, and accept a relatively reduced role for personal relationships and community. In our sample the UK, and US represent individualistic cultural contexts. By contrast, collectivist societies highlight the interdependence between individuals across multitier relationships with an emphasis on one's proscribed duties and responsibilities. Our Macedonian, Mexican and Taiwanese samples represent this view. In addition to individualist and collectivist differences it is also important to note that these cultural contexts differ in religious perspectives. Within our sample there are significant numbers of Christians (Orthodox, Roman Catholic and Protestant), Muslims (primarily Sunni), and adolescents socialized in Eastern traditions.

Given this range of cultural organizations and religious traditions, our interest is whether adolescents apply the virtues in similar ways when assessed using a measure derived from interviews and focus groups with US adolescents. Of specific interest is whether adolescents are more or less similar in making action choices and selecting justifications for action when they interpret and respond to virtue-based moral dilemmas; whether age trends are evident across adolescence; and, whether dilemma content moderates response patterns across cultural contexts. The potential role of cultural context as a moderator of how adolescents apply the virtues follows from the ways in which the social environment supports the development of social/moral understanding. Adolescent development occurs within an ever expanding social context as

adolescents explore their roles within the broader community (e.g., Elder, 2008). As adolescents come to understand their social context, collectivist and individualistic cultures may differentially prioritize the relative importance of friends, family and persons of authority when applying virtue concepts. For instance, the heightened responsibility one attributes to family expectations in collectivist cultures may influence the choices one makes when pressure from friends conflicts with these expectations. Similarly, how one weighs the importance of being forthright to a teacher is potentially different in cultures prioritizing authority to social roles. In our view, culture may alter the calculus of how adolescents weigh the claims of different individuals in their social space both in the choices they make and the justifications for their action.

Measuring the application of virtue concepts

This study is informed in part, by a relatively new class of measurements attending to the cognitive features of moral judgments labeled “intermediate concepts” (Rest, Narvaez, Thoma, & Bebeau, 1999). Intermediate concepts were first identified in relation to the professions (e.g. dentistry) as a missing level of analysis between the highly abstract moral stages (e.g., Kohlberg, 1969) and surface-level professional codes-of-conduct (Bebeau & Thoma, 1999). Researchers in the professions noted that Kohlbergian-based measures, while generally informative in identifying overall moral decision-making strategies of students, were less helpful in assessing student performance on the ethical concepts faced by professionals in the context of their daily practice (Strike, 1982). These researchers requested measures designed to assess reasoning about ethical concepts such as informed consent, privacy, and professional autonomy, and preferred that the evaluation of performance reflect professional standards. To address these preferences, intermediate concept methods were developed as a measurement system designed to

assess targeted concepts as nested within the professional context. These measurement systems presume that the intermediate concept is understood by the individual based on his/her moral judgment processes interacting with contextual factors such as training, experience, precedent, and cultural definitions (Rest et al., 1999).

The measure of moral thinking used in this study follows an intermediate concept approach but expands the focus from the professions to adolescent populations by attending to moral issues embedded within a phase in the life-span rather than a professional setting (Thoma, Derryberry & Crowson, 2013). The focus on adolescence was motivated by both practical and theoretical reasons. Practically, a majority of character education initiatives target adolescent populations and there is a need for theoretically grounded assessments to evaluate these programs (Walker & Thoma, 2018). More importantly a measure targeting adolescence attends to a developmental phase in which the underlying skills used to recognize and attend to moral phenomena goes through marked changes. One notes, for instance, a rapid advance in the adolescent's perspective taking of others as individuals and later as members of social groups (Selman, 1980). Additionally, more traditional measures of moral judgment development (Rest, et al. 1999) describe the adolescent years as the transition from a self-focused view of cooperation to a more system-wide perspective emphasizing norms, social roles and laws. Similarly, virtue theorists identify adolescence as a time in which the habits of childhood become internalized and understood as virtue concepts with a corresponding understanding of how virtues are best expressed (Kristjánsson, 2015). Common to these perspectives is the view that adolescence is characterized as a shift from a self-focused conception of moral concepts to a society-wide perspective.

Is there evidence for a common understanding of the virtues? Evidence that descriptions of the virtues used in the US may have utility in other cultural settings can be found in both our focus on the cognitive aspects of moral thinking as well as empirical assessments of the defining features of virtues. Contemporary empirical support for the cross-cultural utility of the virtues is derived from at least two primary sources: content analyses indicating that the virtues can be identified in various cultural definitions of the good, and research using Peterson and Seligman's Virtue in Action measure (2004). These two supporting areas are derived from many of the same sources and are found under the umbrella of Positive Psychology in general and Martin Seligman's work in particular. In laying the groundwork for the VIA, Seligman and colleagues (Dahlsgaard, Peterson & Seligman, 2005) explored value statements embedded within three clusters of philosophical and religious traditions including Western (subsuming Western philosophical traditions, and Judeo/Christian/Islamic religious traditions), South Asian perspectives (including Buddhism and Hinduism) and Chinese traditions (including Confucianism and Taoism). Through this content analysis, the authors proposed 6 core virtues: courage, justice, humanity, temperance, wisdom and transcendence that could be reliably identified across these different cultures and traditions. Finding that core virtues have generalized utility supported the development and proliferation of the VIA, an objective assessment of one's endorsement of the 24 items reflecting character strengths and virtues (although see recent modifications to the VIA in McGrath, 2015). In a recent summary of the literature using the VIA, Niemiec, (2013) reported similarities in virtue endorsements across 54 nations encompassing both collective and individualistic cultural organizations.

Is there evidence for a common understanding of moral concepts? A similar pattern is noted in the literature on the cognitive aspects of moral judgement development. In both the

Kohlbergian (Colby & Kohlberg, 1987) and Neo-Kohlbergian (Rest, et al., 1999) perspectives, evidence for cross-cultural similarities in the understanding of moral dilemmas and the associated progression of stages or schemas with age and education is sufficient to maintain the proposition of a universal understanding of cooperation and fairness (Gibbs, Basinger, Grime, and Snarey, 2007; Moon, 1986, Gielen, 1996). It appears that across a majority of cultures representing both individualistic and collectivistic social organizations, there is evidence supporting a developmental progression consistent with theory during the adolescent years. When variations to this overall pattern exist they are likely to be noted in adult populations and specifically in the production of post-conventional reasoning (Snarey & Keljo, 1991; however see Rest, et al., 1999). In addition, studies framed by Social Domain Theory suggest that individuals across and within cultures reliably divide the social domain into moral, conventional and personal domains (Turiel, 2002). Turiel further notes that although the content of these domains may vary across contexts, the orientation to other's rights, traditions and personal choice remains a salient cross-cultural feature of social reasoning.

Overall, therefore, there is evidence that moral judgment development and the virtues have some currency across cultures with distinct histories and traditions. However, less well understood is how virtue concepts are applied in context. Although individuals in different cultural settings may understand the concept of courage in similar ways, differences may still exist in how the targeted virtue ought to be expressed within a peer context. Similarly, we know little about how the application of virtue concepts changes over time and in various settings.

Method

Samples and settings. This secondary analysis used data sets that were all part of five research projects completed in the past 10 years. From the Balkans are two samples from

Macedonia including 739 adolescents. Each sample was collected in the capitol city of Skopje and represented students from a gymnasium (N= 389, 61%), a technical high school (N= 249, 31%) and from college (N= 126). Adolescents were sampled by grade which typically corresponds to a particular age. Early high school was represented by N= 412 15 year-olds of which 62% were girls. Late high school included N = 223 17 year-olds equally divided by gender (N= 112 girls). The college student group (N= 104) represented an age range from 19-21 with a modal age of 20. We used the modal age to represent this group. Male college students were overrepresented (N= 73 or 66% of the sample). Participants in the samples were reported to be overwhelmingly Orthodox Christians and members of the majority Macedonian ethnic group

Two urban schools in Mexico provided participants (N= 315) from two age groups: a younger subset (age 12 (N= 110; and age 13, N= 81) and a middle adolescent sample (ages 14, N=78 and 15; N= 315); females were overly represented across both schools (N= 216, 67%). The two schools (one public and the other a private religious school) were situated in a central Mexican city. The public school served a lower socioeconomic section of the city and the other, a Catholic school drew students citywide suggesting some diversity in SES in the total sample. Regardless of school type the majority of students (N= 242, 75%) viewed themselves as practicing Catholics. Ethnic information was not collected.

The Taiwanese sample (N= 1435) was comprised of 21 public and private, urban and rural schools from all of the country's educational regions. The sample focused on 13-15 year olds and included N= 292, 13 year-olds (N = 180 females, 62%), 929, 14 year-olds (N = 460 females, 49%), and N= 214, 15 year-olds (N = 89 females, 42 % females). Participants reported that 39% were practicing their religious beliefs, which included Taoism (35%) and Buddhism (18 %) as

the largest categories. Twenty-nine percent labeled themselves atheist. Participants described themselves as Hokkien (81%), Hakka (15%), Mainland Chinese /new arrivals (7%) and Aborigine (3%), which closely corresponds to the ethnic background of Taiwanese citizens (World Atlas, 2018).

From the United Kingdom, there are three samples: the first is a sample of 14 & 15 year-olds (n=3455) from 39 schools representing England, Scotland, Wales and Northern Ireland and included a mix of private, public, religious and secular schools. In this sample there were N=88, 13 year-olds (59% female), N = 2123, 14 year olds (51 % female) and N = 1207, 15 year olds (52 % female). The majority of this sample did not practice a religion (n=2573, 70 %) and most listed themselves as either Christian (35%) or atheist (non-believer) (34%). The second sample (N= 2350) was similar to sample 1 but included a wider range of ages (12 – 17 years old) and a reduced set of schools. There were N= 852, 12 year-olds (48% female); N=581, 13 year-olds (51% female); N= 384, 14 year-olds (46% female); N= 278, 15 year-olds (44% female); N= 189, 16 year-olds (50% female); and N = 79, 17 year-olds (51% female). Consistent with the first UK sample the majority of participants did not practice any religion. When asked to identify their belief's most claimed to be Atheist (31%), Christian (31%), or Muslim (11%). The final UK sample was obtained from a single school in a large urban area within the Midlands region. The Midlands school drew students from across the city, and tried to reflect the demographics of the area. Two age groups were sampled: N= 169, 12 year-olds (44% female); and N=136, 17 year-olds (60%, female). No religious practice or affiliation information was collected. Overall, the ethnic makeup of the sample (85% white, 5% British Asian, 2% Chinese/other Asian, 1%, Black Caribbean, African, or other Black) reflect UK norms (World Atlas, 2016).

From the US, there are N= 610 adolescents from three samples: a multi-age adolescent sample (N= 156); a high achieving private high school sample (N= 44), and a late adolescent sample (N= 420). The multi-aged adolescent sample included students from two high schools in the Southeast US. One was located in a rural area and the other in an urban setting. Adolescents in this sample represented the typical age range found in US high school (Ns= 29, 56, 36 and 25 for 14 through 17 year-olds). This sample included slightly more males (N= 87 or 56% of the sample). The private school sample was drawn from an urban setting in the Southwest US (Ns= 8, 14, 12, and 10 for 14 through 17 year –olds) and was equally divided by gender. The late adolescent sample was collected in a freshman course within a university setting during the first month of the fall semester. Students were distributed in four age ranges with Ns of 10, 65, 171 and 184 for students aged 17 through 20 respectively. Women were overrepresented in the pool of participants and the resulting sample (N= 339 or 77%). Participants in the college sample described their ethnicity as 78% White, 18% African American, and 3% Hispanic. Religious affiliation was not collected.

This project involved secondary analysis of existing data sets from sources that were reviewed by appropriate ethics boards of the host institutions. As such, this study was deemed exempt by the Institutional Review Board at The University of Alabama (EX-18-CM-046).

Table 1 provides an overview of the samples and associated Ns in the various age groups.

Inspection of Table 1 highlights different interests driving the various studies. The US and Macedonian sample focused on later adolescence whereas the UK researchers sampled the range of adolescents throughout the adolescent years. Researchers in Taiwan and Mexico were primarily interested in younger adolescents. These different interests represent, in part, variation in when each culture emphasizes character education in the schools.

Measures. The primary measure for the study is the Adolescent ICM (AD-ICM). This measure presents adolescents with real-life situations that have been designed to highlight the cognitive aspects of virtue applications. For example, in the honesty story a student is upset that she has not been able to appropriately study for an exam. Another male friend has already completed the exam in an earlier session and offers to help identify the material covered on the exam. The protagonist in the story - also a friend - observes the exchange and is shocked because all students who took the test in the earlier session have pledged not to say anything about the test and its contents. The protagonist wonders what to do. After reading this and the other stories, the respondent is asked to rate individual action choices as more or less appropriate and then rank the four most appropriate action choices. Following consideration of the actions choices, the participant repeats the process by rating and ranking the best justifications for an action. Additionally, participants are asked to identify the most inappropriate action choices and justifications.

High scores on the measure represent a response pattern in which the appropriate and inappropriate choices and justifications match experts' interpretation of the items as acceptable, unacceptable or neutral. Low scores, by contrast, are the result of the participant ranking items as appropriate that the experts judge to be unacceptable as well as the reverse. As such, scores on the AD-ICM represent the degree to which the participant applies the targeted virtue in a manner consistent with the prevailing view of what constitutes appropriate applications of the virtues. It is important to note that the ICM approach does not assume a single best choice or justification. Instead, it presupposes that there are multiple appropriate (and inappropriate) choices that as a set can be distinguished from less appropriate possibilities (Bebeau & Thoma, 1999). Current empirical evidence suggests that the AD-ICM is a reliable measure that produces

trends consistent with theoretical expectations for an ICM measure (Thoma, et al., 2013; Walker, et al., 2017).

Two versions of the AD-ICM are summarized in this study: the original Adolescent ICM (Thoma, et al., 2013) and a short form developed for the UK, Taiwan and Mexican studies (Walker, et al., 2017). Both versions use the same rating and ranking task and differ only in the number of stories. Action choice and justification items were written to provide a minimal framing of the key rationale for the choice/justification. This approach follows methodological observations in the development of objective measures of moral judgment development (Rest, 1979). In this literature it was found that more stable estimates of moral judgments were obtained when the individual is asked to judge items with limited information that is just sufficient to trigger an interpretation of the item's meaning rather than asking participants to endorse developed arguments (see also Rest et al., 1999).

The AD-ICM in its traditional form includes seven stories and takes approximately 45 minutes to complete. The short version includes three stories derived from the original seven and takes about 20 minutes to complete. The overlapping stories are the same in all respects with two exceptions: the courage story used in the UK, Mexican and Taiwan samples uses an alternative setting while keeping constant the central issue and items. Second, some basic wording changes were made to reflect culturally specific terms and informal language.

The steps to insure cross-cultural equivalence were similar for each of the non-US samples. Following translation from the original US version, the revised measure was piloted with the intended population to assess familiarity with the situations described in the stories as well as their relevance to the adolescent experience. Interestingly, in each case the research team found that the AD-ICM stories captured situations and issues that adolescents found to be relevant and

realistic. Beyond the wording changes to better reflect local usage, researchers reported the measure development process to be straightforward and major changes to stories and items were not needed. The specific stories and associated virtue concepts are provided in Table 2.

The scoring of participants' item rating and rankings is accomplished by attending to the degree of fit to a key developed by an expert panel. The panel evaluated each story and associated items, and then judged each item as acceptable, unacceptable or neutral. Specifically, items were rated as acceptable if the experts agreed that the ideas expressed by the item could be justified. Similarly, unacceptable items represented ideas that the experts agreed were problematic. Experts were defined as graduate students in adolescent psychology with training in moral psychology (see Thoma, et al., 2013, for a discussion of the use of this group as an expert panel).

In addition to an overall score, this scoring process yields sub-scores for "unacceptable choices" (the degree to which the participant identifies action choices and justifications the experts see as unacceptable) and "acceptable choices" (the degree to which the participant identifies action choices and justifications the experts also see as acceptable). Additionally, the measure provides separate scores for action choices and justifications. Thus, in addition to the overall AD-ICM score the measure generates two sets of sub scores: acceptability (acceptable and unacceptable choices) and choice type (action choices and justifications). Scores for the total score as well as the sub-scores are represented as percentages so for example a score of .60 indicates that 60% of the participant choices match the expert derived key.

Reliability of the AD-ICM. There are at least three reasons why one does not expect high internal consistency estimates using the AD-ICM. First and despite the involved rating and ranking tasks associated with each dilemma, the only statistically independent unit of measure is

the story and as such, the long form of the survey equates to a 7-item scale and the short form a 3-item scale. Typically, scales of this length do not achieve high internal consistency estimates. Second, some of the samples had narrow age-ranges and potentially suffered from restriction of range. Such restrictions of range reduce alpha estimates (Thorndike, 1988). Finally, the assessed virtues are broad constructs and are measured within different contexts. The degree to which participants attend to different features of the virtue and context across the stories will reduce the estimate of internal consistency.

As expected, there was a relationship between number of stories and obtained alpha estimates (Alpha US = .79; Alpha Macedonia = .69; Alpha UK = .53, Alpha Taiwan = .51, Alpha Mexico = .55). To provide a common metric we scaled each short-form estimate to the 7-item equivalent and find a consistent pattern of alphas in the mid-70s. Not surprisingly, the US sample had both the broadest age-range and highest internal consistency estimate.

Procedures. In addition to the primary AD-ICM measure, each study collected gender and age information. Additional measures were often included to serve the specific needs of the original study. These additional assessments included other inventories and extensive requests for information about the participant such as about family characteristics, the parents' educational backgrounds, personal interests and activities.

Two main methods were used to collect participant responses: paper and pencil and on-line web-based questionnaires. Within some of the cultural settings both systems were employed (U.S. and U.K. –about half of the participants were assessed using a paper and pencil approach). The Macedonian sample used paper and pencil only whereas the Taiwan and Mexican samples used an online system within a classroom setting. The data were scored in a single location using the same key and algorithm to insure a standardized treatment of the data.

Data analysis plan. To address performance on the AD-ICM across and within the different cultural contexts we adopted a two-step strategy. First, using ANOVA, we assess the overall trends across the samples noting whether AD-ICM scores increase with age, differ by story, differ by subscales (i.e., action choices and justifications, acceptable vs. unacceptable responses) and differ by gender. The analyses identified to assess cultural context differences were less straightforward because of the non-overlapping age groupings across samples. When age was not included in the analysis then culture was included as a between-subjects factor and used to test whether cultural context moderates the effect of interest. For example, in the assessment of story differences across cultural contexts, we used a one within (story) and one between (cultural contexts) repeated measures ANOVA. However, age was of importance in a majority of analyses. In these cases, we assess the effect of interest using the same analyses used in the secondary analysis but applied within each cultural context. To compare the findings across the five cultural contexts we attend to effect sizes and their confidence interval estimates (Cumming & Finch, 2001). If effect sizes are different across the cultural contexts (i.e., fall outside the confidence intervals) we suggest that context moderates the overall findings and adolescences in these contexts perform differently on the AD-ICM. Taken together, the hybrid analytic approach adopted in the current study pairs a traditional secondary analysis for the overall assessments with a meta-analytic component for the nested results by culture.

Results

Results are ordered by first attending to the overall AD-ICM scores followed by story differences, acceptability subscales (acceptable and unacceptable choices/justifications), type of judgment (action choice vs. justifications) and gender. In each section, the omnibus test is

followed by the individual assessments within the five cultural contexts. The descriptive statistics associated with these analyses are presented in the supplemental analyses.

Age/educational Trends

Older adolescents achieve higher total AD-ICM scores than their younger peers ($F(8,9111) = 23.508, p < .05; \eta_p^2 = .02$). Post hoc comparisons using corrections for unequal variances suggested that the overall AD-ICM means follow a curvilinear pattern where scores decline in early adolescence reaching a statistically significant lowest point at age 15 when scores begin to increase. In older adolescent populations, the difference by age was more pronounced. Specifically, 18 year olds were higher than the younger adolescents and 19 and 20 year olds were higher than all other groups. A test of the nonlinear features of the relationship between age and AD-ICM total scores confirmed a quadratic relationship ($F(2,9109) = 57.11, p < .05, \eta_p^2 = .02$). Taken together these findings are consistent with the view that the ability to apply the virtues is developmental across adolescence.

These overall trends mask some variation in average total AD-ICM scores across groups. Figure 1 provides age trends for each cultural context. Although the comparisons are not fully overlapping, the data indicate evidence for upward change, particularly for the samples which provided an extended age-range. In the group assessing a broad age-range are the US, UK and Macedonia. Each of these samples indicated significant age differences (US: $F(6, 558) = 11.03, p < .05; \eta_p^2 = .11$; Macedonia: $F(2, 734) = 70.97, p < .05, \eta_p^2 = .16$; UK: $F(5, 5834) = 8.03; \eta_p^2 = .01$). Inspection of the confidence intervals indicates that US and Macedonian age effects were greater than the UK effect. The Taiwanese and Mexican samples are noticeable for higher scores across the sampled age ranges but little evidence of growth. The lack of age trends for these two contexts is not surprising given the focus on younger adolescents and narrow age-ranges.

Age trends by story and cultural context. To further explore age trends, we assessed cross-sectional patterns by story. Of particular interest was the dip in scores during middle adolescence, and whether story content moderated the age/education effects by cultural context. To insure comparability across cultural contexts, we focused on the three stories common to all samples. A repeated measures ANOVA on AD-ICM scores with cultural grouping as the between subjects factor and story as the within subject factor resulted in main effects for story ($F(1.924, 102.82) = 844.95, p < .05, \eta^2 = .08$), cultural grouping ($F(4, 9168) = 149.16, p < .05, \eta^2 = .06$), which was conditioned by a group by story interaction ($F(7.70, 14,00) = 115.03, p < .05, \eta^2 = .05$). Inspection of the means indicated that the David story (self-discipline) was associated with the highest scores across cultural contexts followed by the Janie story (Courage). The most difficult story for participants was Danielle in which honesty is pitted against friendship and loyalty considerations. As the interaction of group and story suggests the pattern in mean story scores varied by cultural groupings. To interpret this interaction, Figure 2 presents the story means by cultural groupings. This figure indicates that in three of the cultural groupings (US, UK and Macedonia), the story differences were most pronounced. Noticeably different from these three were the Taiwanese and Mexican samples where performance on the Danielle, David and Janie story were much more uniform. Additionally, it appears that the higher total AD-ICM scores for the Mexican and Taiwanese samples noted in the previous section can be attributed to more consistent performance across stories and particularly on the Danielle story addressing honesty.

To further explore the interaction between cultural context and story on AD-ICM scores we wondered whether differences in age can account for the story differences noted above given the differences in average age associated with each cultural group. To address this question we

analyzed the story data limiting age to adolescents under 16 in order to assess whether the Mexican and Taiwanese participants performed differently because they were on average the youngest participants. The results of this analysis did not differ from the overall patterns and confirmed that the Taiwanese and Mexican participants did not have the same difficulty with the honesty (Danielle) story that was evident in the other three groups.

Age trends on the subscales.

In addition to the overall scores, the AD-ICM provides subscales, which can be used to address specific differences in the ability to identify choice type (justifications from action choices) as well as acceptability (acceptable choices and justifications from inappropriate ones). Both of these contrasts have been noted using ICM measures in the professions (Bebeau & Thoma, 1999, Turner, 2010) as well as when using the AD-ICM (Thoma, et al., 2013; Walker, et al., 2017). The AD-ICM subscales were assessed by age, and by cultural context using the same approach applied to the assessments of story differences.

Action choices vs. justifications. On the relative ability to identify justifications and action choices, the overall analysis indicated that justification scores lagged behind action choice scores in younger adolescent populations. However, by age 17 the two scales converged. Specifically, a repeated-measures ANOVA on subscale scores with action choice and justification rankings as the within subjects factor and age as the between subjects factor found an effect for type of judgment ($F(1, 9101) = 23.78, p < .05, \eta^2 = .01$) and age ($F(8, 9101) = 23.49, p < .05, \eta^2 = .02$). These main effects were conditioned by a subscale by age interaction ($F(8, 9101) = 5.92, p < .05, \eta^2 = .01$). As Figure 3 indicates, adolescents are better at identifying action choices than the justifications for action, particularly in early adolescence.

In addition to the overall analyses, a repeated measures ANOVA with choice type as the within, and age as the between effect was applied within each cultural grouping. The main effect for decision type was observed in all five of the cultural contexts ($\eta_p^2 = .01, .03, .02, .01, .02$ for the US, Macedonia, UK, Taiwan and Mexico respectively). Based on the confidence intervals the individual effects did not vary across cultural contexts. Interactions with the age variable were found in only the US and UK samples ($\eta_p^2 = .06, .02$). Overall it appears that the difference in performance in identifying appropriate justifications and action choices is more evident in younger adolescent populations and may decline with age.

Acceptability. We also wondered whether adolescents identified appropriate choices and justifications differently from the selection of inappropriate choices. Earlier reports in the US and UK (Thoma, et al., 2013; Walker, et al., 2017) indicate that these different choices may actually describe different applications of the virtues and are not simply the mirror image of the other. Specifically, these earlier studies found consistently higher scores for acceptable choices when compared to unacceptable scores. To assess the overall differences in acceptability scores, we applied a repeated-measures ANOVA with acceptability score as the within subject factor (acceptable and unacceptable) and age as the between subjects factor. Our overall results confirm a moderate main effect favoring the appropriate choices ($F(1, 9103) = 432.58, p < .05, \eta_p^2 = .05$) which was not conditioned by an interaction with age ($F(8, 9103) = 1.43, p > .05, \eta_p^2 = .00$). Similar analyses assessed within cultural groupings indicated main effects for acceptability in all five of the cultural groupings with effect sizes ranging from moderate ($\eta_p^2 = .07, .17, .17$, for the UK, , US and Taiwan samples respectively) to large ($\eta_p^2 = .23, .24$ for Macedonian and Mexican samples respectively). Only the UK effect size was different (smaller) from the other four cultural contexts.

We further explored acceptability judgments at the story level using a repeated measures ANOVA, with acceptability judgments and story as within subject factors and cultural context as the between subject factors. As expected from previous analyses, we found statistically significant effects for story and appropriate/inappropriate choices. In addition we found a moderate interaction effect between story and acceptability ($F(1.96, 17996.95) = 1416.26, p < .05, \eta_p^2 = .13$) and a three-way interaction between story, cultural context and acceptability ($F(7.85, 17996.95) = 39.79, p < .05, \eta_p^2 = .02$). This finding suggests that the difference in identifying acceptable and unacceptable choices varies by story and to a lesser degree by context.

When assessed within each cultural group using age as the between-subject factor we found similar effects for the story by acceptability interaction ($\eta_p^2 = .14, .23, .31, .43, .40$ for the US, Macedonian, UK, Taiwan, and Mexico samples respectively). The US and Macedonian acceptability effect sizes were statistically lower than those found in the other three cultural contexts. Three-way interactions with age were found only in the US and Macedonian samples and these effects were small. Taken together the overall main effect for acceptability found at the summary level masks differences by the specific virtue and context. In general, our findings indicate an advantage in identifying appropriate choices and justifications as compared to the inappropriate choices that persists across the adolescent years, but this advantage varies in magnitude by the specific virtue and cultural context.

Gender differences.

Females consistently achieve higher scores on the AD-ICM (Thoma, et al., 2013; Walker, et al., 2018) and on other ICM profession-specific measures (e.g., Bebeau & Thoma, 1999). Across cultural groups, we also found gender difference favoring females ($F(1, 9072) = 56.47, p < .05, \eta_p^2 = .04$) which was not conditioned by an interaction with age ($F(8, 9072) = .22, p >$

.05, $\eta^2 = .00$). A gender effect was also observed within the five cultural groups and the magnitude of the effect did not differ from each other ($\eta^2 = .08, .06, .02, .06, .09$ for the US, Macedonian, UK, Taiwan and Mexican participants respectively).

To assess whether gender moderated AD-ICM findings within the cultural groupings, we used repeated measures ANOVAs on the within-subject effects of story, type of decision, and the ability to identify acceptable and unacceptable choices by gender and age. As with the overall analyses there were few gender moderated effects observed within groups and when they occurred the effects accounted for little variance. For story differences by gender, only the Macedonian and Taiwanese results indicated statistically significant results and these interactions presented a pattern of means that were very similar to the other cultural groups and differed only by degree. Similarly, within group comparisons between age and types of choices yielded one gender by choice interaction associated with the Taiwanese sample. Again, the nature of the interaction was not distinguishable from the other cultural groupings. Finally, and perhaps most interestingly, when gender was assessed with age and acceptability within the US and Macedonian samples, we found evidence of a moderating role for gender. In both samples, the nature of the interaction indicated a diminishing gender difference with age that was most evident in 17-year-old groups and above. Overall, the most striking gender effect was the overall difference favoring women on AD-ICM scores, which was also found within each cultural context.

Discussion

Using the newly developed AD-ICM, the current study provides a secondary analysis of adolescent data drawn from 5 distinct cultural contexts. The goal of these analyses was to examine similarities and differences in moral decision-making in the application of the virtues

with a particular interest in developmental change, both generally and also on particular virtues and contexts. Additionally, gender was of interest given the well-established female advantage on intermediate concept measures of moral thinking (e.g., Bebeau & Thoma, 1999). To address these questions, we compared samples drawn from cultures that differ on an individualistic to collectivist dimension as well as samples representing various religious perspectives including: predominantly Christian (US, UK (majority Protestant) and Mexico (majority Catholic), a mixture of Christian (Orthodox) and Muslim (UK and Macedonia) and Eastern (Taiwan). The results from this study suggest both similarities and differences in adolescents' application of the virtues across these cultural contexts. These similarities are particularly interesting in that the stories and items used in this study were developed within a US adolescent population. It may be that the role of context is less central when the intermediate concepts are based around the virtues because these concepts are recognizable within a variety of Eastern and Western ethical traditions (Shanahan & Wang, 2003) and relate to situations likely to be experienced by most youth. To highlight this point, it is interesting to note that the process by which the various stories and items were adapted to the different cultural contexts was relatively straightforward and typically required only surface-level word changes (e.g., changing "great!" to "brilliant!"). The ease with which these situations and items were modified suggests that the virtues and the assessment context support the utility of the measure beyond the US context.

These similarities in responses notwithstanding, it may be that our observed findings mask true differences between the sampled cultural contexts which we are unable to directly assess. For instance, one could argue that the AD-ICM provides only proxies for virtue applications and not the reasoning process underlying these choices. Thus, while the observed choices may appear similar across cultural contexts the underlying considerations informing

these choices as well as their cognitive and affective associations are actually quite different (e.g., Haste, 2013). Although we agree with this concern, we note that the AD-ICM requires participants to rate and rank action choices and justifications nested within a specific social context. As such, our measure provides multiple indicators of how the applications of virtues are understood and ought to be more sensitive to the possibility of false similarities than the traditional inventories targeted for criticism by cultural psychologists (e.g., Ratner & Hui, 2003). Indeed, our attention to story and subscale scores was motivated by the concern that differences between contexts might be more subtle and tied to particular types of decisions. Overall we suggest that our findings are consistent with the view that there are similarities across cultural contexts. However, we recognize the need for future studies exploring how virtues function within social structures and belief systems to more fully assess the generalizability of virtue applications across cultural contexts.

Age-related Trends

Evidence for age trends on AD-ICM scores was found in the US, UK and Macedonian samples. The magnitude of these differences varied from moderate to small effects. Common to the samples demonstrating age trends was a broad sampling of adolescents that spanned the adolescents period whereas the two samples that had younger and more homogeneous samples did not provide evidence for change. Age-related trends were also evident on the subscales attending to judgments about action choices and justification as well as on the acceptability scales.

Within these general trends, however there is some variation in average scores across groups. Specifically, Taiwanese and Mexican young adolescents scored higher than their same aged peers in the US, Macedonia and the UK. The higher scores of these two groups were

associated with less difficulty on the honesty dilemma and on identifying appropriate action choices. It is unclear whether these mean differences reflect real performance variations or issues related to the relevance of the stories in non-US contexts. However, it is interesting to note that the two groups associated with the highest scores are plausibly the most affected by story relevance and translation issues and thus it is unlikely that measurement weaknesses are major contributors to these group findings. Whether or not these mean differences represent measurement concerns it seems reasonable to suggest that US populations do not have an advantage on the measure despite having been the context emphasized in the development and validation process.

Differences by Story and Context

The three stories in common to the five groups were assessed for differences both across and within cultural contexts. Taken together, statistically significant differences were noted between the dilemmas (Self-discipline > Courage > Honesty). However, there were context by story differences. Particularly noteworthy was the relatively consistent performance across the three stories found in the Mexico and Taiwan samples. By contrast, the other more Western and European samples presented variation across stories and especially a noticeable lower score on the honesty story. As mentioned previously, in this story participants must weigh a pledge to a teacher not to divulge an exam's content against reporting a friend. Particularly salient in the dilemma is the potential damage to the central character's friendships and a friend's academic standing. Both male and female participants in the Western and European samples were more unwilling to set aside friendship considerations. It may be that the seemingly reduced focus on friendships in the Taiwanese and Mexican samples reflects the relative priority placed on 'the collective' and roles within the group (e.g., the teacher) rather than individual relationships. In

support of this view, there is evidence both within the literature on cognitive moral judgments (Snarey, 1985) and in related social cognitive constructs (e.g., empathy) highlighting differences in performance due to collectivist or individualist cultural contexts (Chopik, O'Brien, & Konrath, 2017). Although perceived as a collectivist culture much like Taiwan and Mexico, the Macedonia findings followed the pattern found in the US and UK samples. This seemingly contradictory result may reflect other cultural differences beyond the more general collectivist/individualist dimension. For instance, cultural maps based on multiple cultural value dimensions locate Macedonia in closer proximity to European countries and the UK (Inglehart, 2003). Thus, it is plausible that the role of friendships in Macedonian adolescents is more closely aligned with Western views.

Subscale differences by age and group. Our focus on the AD-ICM subscales helped to clarify more subtle aspects of the developmental process. On the contrast between justifications and action choices, it appears that across the adolescent years these two judgments become more coordinated and by age 17 differences cease to be meaningful. Indeed, it appears that late adolescence is a period in which similarities both across and within cultural groups become more noticeable. Our ability to state this definitively is weakened by the limited age-ranges of two of our more collectivist samples. However, in those samples which include older adolescents, we find action choice, justifications, and average AD-ICM scores each converging in the older age groups. Additionally, in two of these groups gender differences become less pronounced in older adolescents. Why, then, do AD-ICM scores tend to converge across adolescence?

A plausible explanation follows from the observation that adolescence is characterized by normative social and cognitive transitions, which influence how the individual understands the social world. These developmental transitions increasingly orient the adolescent to norms and

values of the larger social environment (Selman, 1980; Rest, et. al., 1999). For instance, we note that the Neo-Kohlbergian literature describes adolescence as a period of transition from a personal interest view of social cooperation (i.e., judgments which emphasizes personal outcomes and relationships) to a maintaining norms perspective (i.e., judgments framed around social norms and laws) (Rest, et al., 1999). This developmental transition can also be understood to describe a shift from a self-focused to a society-wide perspective-taking orientation (e.g., Selman, 1980).

We suspect that the adolescents' growing recognition of societal-wide definitions of cooperation is also reflected in an understanding of the virtues that increasingly aligns with a broader social view exemplified by the AD-ICM experts. That is, the normative shift to a maintaining norms perspective promotes a greater similarity in the understanding of virtue concepts across cultural contexts because of historical cross-influences or similar understandings of a virtuous life. In contrast, personal interest considerations tend to emphasize the unique features of the local context as the individual prioritizes concerns about desirable outcomes for the self, family and friends. We see this interpretation of our findings as generally supportive of theorists who describe the adolescent years as a period of shifting cultural influences. For instance, as Erikson (1968) noted and others have shown (e.g., Selman, 1980), the role of the larger context is especially influential for social development during adolescence but the nature of this influence shifts from important individuals who have direct impact on the adolescent to more society-wide agents. Our view follows this basic proposition and suggests that cultural context effects may not be constant across development particularly during the adolescent years.

Acceptability judgments. Also of interest is the difference in acceptable and unacceptable responses on the AD-ICM across contexts. Unlike the justifications and choice contrasts,

identifying the acceptability of different virtue application choices did not change with age but instead indicated a clear advantage in the ability to identify acceptable choices and justifications over unacceptable selections. This effect was stable and replicated across all cultural groupings.

The clear and persistent differences between acceptability scores may be due to features of the measurement system. In support of this view we note that the measure always solicits the positive choices first. Therefore the AD-ICM procedures may be priming positive applications of the virtues and in so doing disadvantage the selection of unacceptable choices. This methodological consideration notwithstanding, we note that at the very least, knowing what a good choice is does not automatically extend to identifying what is inappropriate. An interpretation consistent with these findings suggests that different cognitive processes and social experiences may contribute to identifying “good” and “bad” choices. Interestingly, within the field of moral psychology there is a tradition of differentiating the ability to attend to actions that are considered “bad”/deficient/rule violations from the more prosocial/good/rule following approaches (Smetana, 2006). Our findings are consistent with this differentiated view.

In addition, differences in identifying the “good” and “bad” may be a result of the ways in which we experience and learn about the virtues. For instance, we wonder whether there is less of a focus on the “bad” when children and youth come to understand the implications of virtue concepts. First, it is likely that we experience good behaviors at a much higher rate than bad behaviors and thus we have more experiences with, and a better understanding of appropriate choices and justifications for action (Walker, et. al., 2017). Similarly, when we are instructed—either informally (e.g., in the family) or formally (e.g., in school) about how one should understand the virtues—the focus of these interactions may emphasize the positive or appropriate. While understandable as an instructional focus, the emphasis on the good requires

the individual to then infer what is inappropriate and poorly conceived. As our data indicate, these inferences appear more difficult. To counter this unevenness, our work suggests that moral education may benefit by expanding its focus to highlight both negative and positive manifestations of virtue concepts in order to provide more guidance in developing an experiential base to support more optimal moral decisions. In support of the benefits of an increased focus on the inappropriate it is interesting to note that the difference in identifying appropriate and inappropriate choices is not found in professions (e.g., Military officers) in which minimizing negative outcomes is a major instructional focus (Turner, 2010).

Gender Differences

Why is it that females are more advanced than their male peers? Early in the construction of the measure it was thought that the female advantage was due to females having a stronger voice in the test development process (Thoma, et. al. 2013). Anecdotally, researchers involved in the construction of the measure recalled more detailed written responses from women and a heightened interest in the topics and issues under discussion. However, in this study we find a consistent pattern of gender differences favoring females across contexts and age groups suggesting a more interesting performance difference by gender. To support this view, we note that our findings are consistent with gender differences on other moral and character measures (e.g., Thoma, 2006; Wang, et al., 2015), and is of the magnitude typically observed in intermediate concept measures in professional populations (e.g., Bebeau and Thoma, 1999).

It should be noted that the well-known female advantage on verbal measures could underlie gender differences on a measure such as the AD-ICM since the task requires reading and ranking items and weighing different choices and justifications. Whether the observed gender difference is specific to social/moral reasoning or the result of some combination of factors including more

general cognitive/verbal skills, we suggest that further attention is warranted on gender's role in the application of virtues in everyday life.

Conclusions

At the beginning of this work we wondered whether there are similarities and differences in the understanding and application of virtue concepts across various cultural contexts. It was our expectation that cross-cultural similarities were due to the overlap in cultural understandings of what constitutes a good moral life, the globalization of western education systems (e.g., Serpell & Hatano, 1997), as well as empirical evidence from moral psychology on cross-cultural features of moral thinking. Our findings are generally consistent with this view but our ability to fully address our interests was hampered by sample characteristics. This limitation constrained our ability to test more subtle questions; most notably the generalizability of a convergence in AD-ICM scores in late adolescence. These limitations notwithstanding, the current study supports the notion that there is utility in viewing the application of virtues across multiple cultural contexts, and supports further work that seeks to test more nuanced cross-cultural claims.

References

- Bebeau, M., & Thoma, S. J. (1999). "Intermediate" concepts and the connection to moral education. *Educational Psychology Review*, 11, 343-360.
- Chopik, W. J., O'Brien, E., & Konrath, S. H. (2017). Differences in empathic concern and perspective taking across 63 countries. *Journal of Cross-Cultural Psychology*, 48(1), 23-38.
- Colby A., & Kohlberg, L. (1987). *The measurement of moral Judgment*. Cambridge: Cambridge University Press.
- Cumming, G. & Finch, S. (2001). A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions, *Educational and Psychological Measurement*, 61, 532-574.
- Dahksgaard, K., Peterson, C., Seligman, M.E.P. (2005). Shared Virtue: The Convergence of Valued Human Strengths Across Culture and History. *Review of General Psychology*, 9, 3 203-213.
- Elder, G. H. (2008). The Life Course as Developmental Theory. *Child Development*, 69, 1, 1-12. doi.org/10.1111/j.1467-8624.1998.tb06128.x
- Erikson, E.H. (1968). *Identity, youth, and crisis*. New York: Norton.
- Gibbs, J.C., Basinger, K.S., Grime, R.L, & Snarey, J.R. (2007). Moral judgment development across cultures: Revisiting Kohlberg's universality claims. *Developmental Review*, 27, 4, 443-500.
- Gielen, U. P. (1996). Moral reasoning in cross-cultural perspective: A review of Kohlbergian research. *World Psychology*, 2, 313-333.

- Haste, H. (2013). Deconstructing the elephant and the flag in the lavatory: promises and problems of moral foundations research. *Journal of Moral Education*, 42, 316-329.
- Hofstede, G. (2001). *Culture's Consequences: Comparing values, behaviors, institutions and organizations across nations*. (Second Edition). Thousand Oaks, California, Sage Publications.
- Inglehart, R. (2003). *Human Values and Social Change: Findings from the Values Surveys*. International Studies in Sociology & Social Anthropology (Book, 89). Leiden, The Netherlands: Brill Academic Publications.
- Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. A. Goslin (Ed.), *Handbook of socialization theory* (pp. 347-480). Chicago: Rand McNally.
- Kristjánsson, K., (2015). *Aristotelian character education*. New York: Routledge
- McGrath, R. E. (2014). Character strengths in 75 nations: An update. *The Journal of Positive Psychology*, doi=10.1080/17439760.2014.888580
- McGrath, R. E. (2015). Integrating psychological and cultural perspectives on virtue: The hierarchical structure of character strengths. *The Journal of Positive Psychology* <http://dx.doi.org/10.1080/17439760.2014.994222>
- Niemiec, R. M. (2013). VIA character strengths: Research and practice (The first 10 years). In H. H. Knoop & A. Delle Fave (Eds.), *Well-being and cultures: Perspectives on positive psychology* (pp. 11-30). New York: Springer.
- Moon, Y. L. (1986). A review of cross-cultural studies on moral judgment development using the Defining Issues Test. *Behavior Science Research*, 20, 147-177.

- Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. Oxford: Oxford University Press.
- Ratner, C. and Hui, L. (2003). Theoretical and methodological problems in cross-cultural psychology. *Journal for the Theory of Social Behaviour*, 33: 67–94. doi:10.1111/1468-5914.00206
- Rest, J., Narvaez, D., Thoma, S. J. & Bebeau, M. J. (2000). A Neo-Kohlbergian Approach to Morality Research. *Journal of Moral Education*, 29, 381-396.
- Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being*. New York: Free Press
- Selman, R. L. (1980). *The growth of interpersonal understanding*. New York: Academic Press.
- Serpell, R. & Hatano, G. (1997). Education, schooling and literacy in a cross-cultural perspective, In J.W. Berry, P.R. Dasen, T.S. Saraswathi (Eds.), *Handbook of Cross-cultural Psychology*, (Vol. 2, pp. 339–376). Basic Processes and Human Development, Allyn and Bacon, Boston
- Shanahan, T. and Wang, R. (2003). *Reason and insight: Western and Eastern perspectives on the pursuit of moral wisdom* (2nd). Belmont, CA: Wadsworth.
- Smetana, J. G. (2006). Social-cognitive domain theory: Consistencies and variations in children's moral and social judgments. In M. Killen & J. Smetana (Eds.), *Handbook of moral development* (pp. 119–153). Mahwah, NJ: Lawrence Erlbaum Associates.
- Snarey, J. R. (1985). Cross-cultural universality of social-moral development: A critical review of Kohlbergian research. *Psychological Bulletin*, 97(2), 202-232.

- Snarey, J.R., & Keljo, K. (1991). The cross-cultural expansion of moral development theory. *Handbook of moral behavior and development* (Vol. 1, pp. 395-424). Hillsdale, NJ: Erlbaum.
- Strike, K. A. (1982). *Educational policy and the just society*. Chicago, Illinois University Press.
- Thoma, S. J. (2006). Research using the Defining Issues Test. In M. Killen & J. G. Smetana (Eds.), *Handbook of Moral Psychology* (pp.67-92). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. .
- Thoma, S. J., Derryberry, P. & Crowson, H. M. (2013). Describing and testing an intermediate concept measure of adolescent moral thinking. *European Journal of Developmental Psychology* (special issue), 10, 239-252.
- Thorndike, R.L. (1988). Reliability. In J. P. Keeves (Ed.). *Educational research, methodology, and measurement: An international handbook*. Oxford, UK: Pergamon Press.
- Turiel, E. (2002). *The culture of morality: Social development, context, and conflict*. New York: Cambridge University Press.
- Turner, M. (2010). *Describing and testing the army leadership ethical reasoning test*. Unpublished dissertation, The University of Alabama.
- Walker, D. I., Thoma, S. J., (2018) Moral and character education, Oxford University Press: *The Oxford Research Encyclopedia of Education*.
- Walker, D.I., Thoma, S. J., Jones, C. & Kristjánsson, K. (2017). Adolescent moral judgement: A study of UK secondary school pupils. *British Educational Research Journal*. DOI: 10.1002/berj.3274
- Wang, J., Hilliard, L.J., Hershberg, R.M., Bowers, E.P., Chase, P.A., Champine, R.B., Buckingham, M.H., Braun, D.A., Gelgoot, E.S., & Lerner, R.M. (2015). Character in

Childhood and early adolescence: models and measurement. *Journal of Moral Education*,
44(2), 165-197.

World Atlas (2018). Retrieved from <https://www.worldatlas.com>.

Table 1.

Age and education distribution of participants by cultural context and age groupings.

	Age									Total
	12	13	14	15	16	17	18	19	20	
Countries										
Macedonia				412		223			104	739
Mexico	111	81	79	44						315
Taiwan		292	929	216						1435
UK	773	670	2507	1486	189	215				6091
US			37	70	48	35	65	171	184	610
Totals	884	1043	3552	2228	237	473	65	171	288	9190

Table 2.

AD-ICM stories and themes.

Primary virtue concept	Story Protagonist	Short form	Summary Description
Fairness	Nikki		Whether or not to fire a friend who is the weakest worker.
Responsibility	Chris		Whether or not to avoid a commitment and go out with friends.
Loyalty	Molly		Supporting someone when it is difficult to do so.
Self-discipline	David	X	Continuing to prepare for an important exam or go on a class trip.
Honesty	Danielle	X	What to do when friends cheat in school.
Courage	Janie	X	Sticking to one's beliefs at the cost of recognition and success.
Respect/Loyalty	John		Fulfilling one's personal commitment to a mentor when abandoning him might result in greater personal gain.

Note: "X": indicates that the story is included in the short form.

Figure 1. Age trends on AD-ICM total scores by cultural contexts.

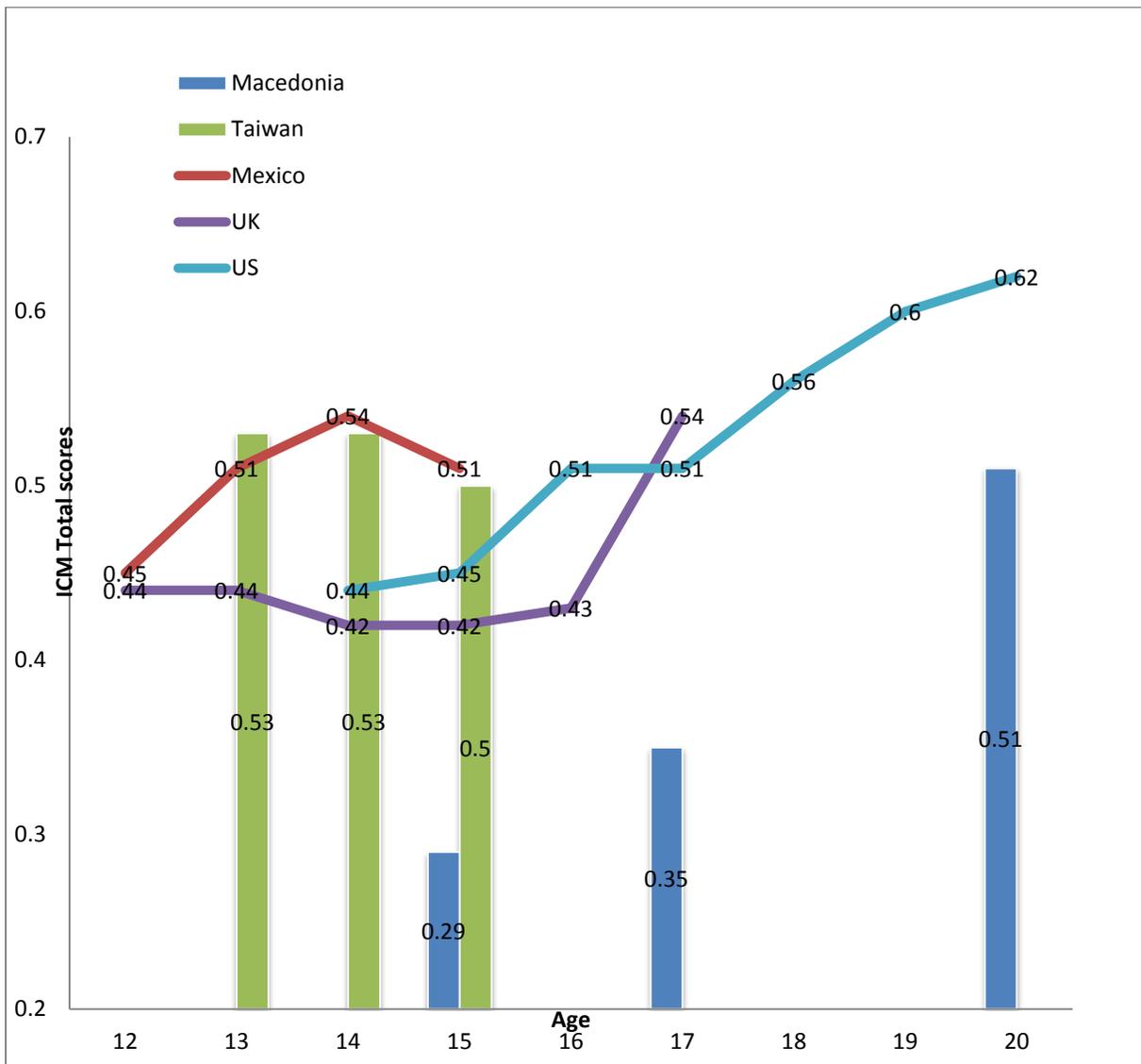


Figure 2. Story Scores by cultural contexts.

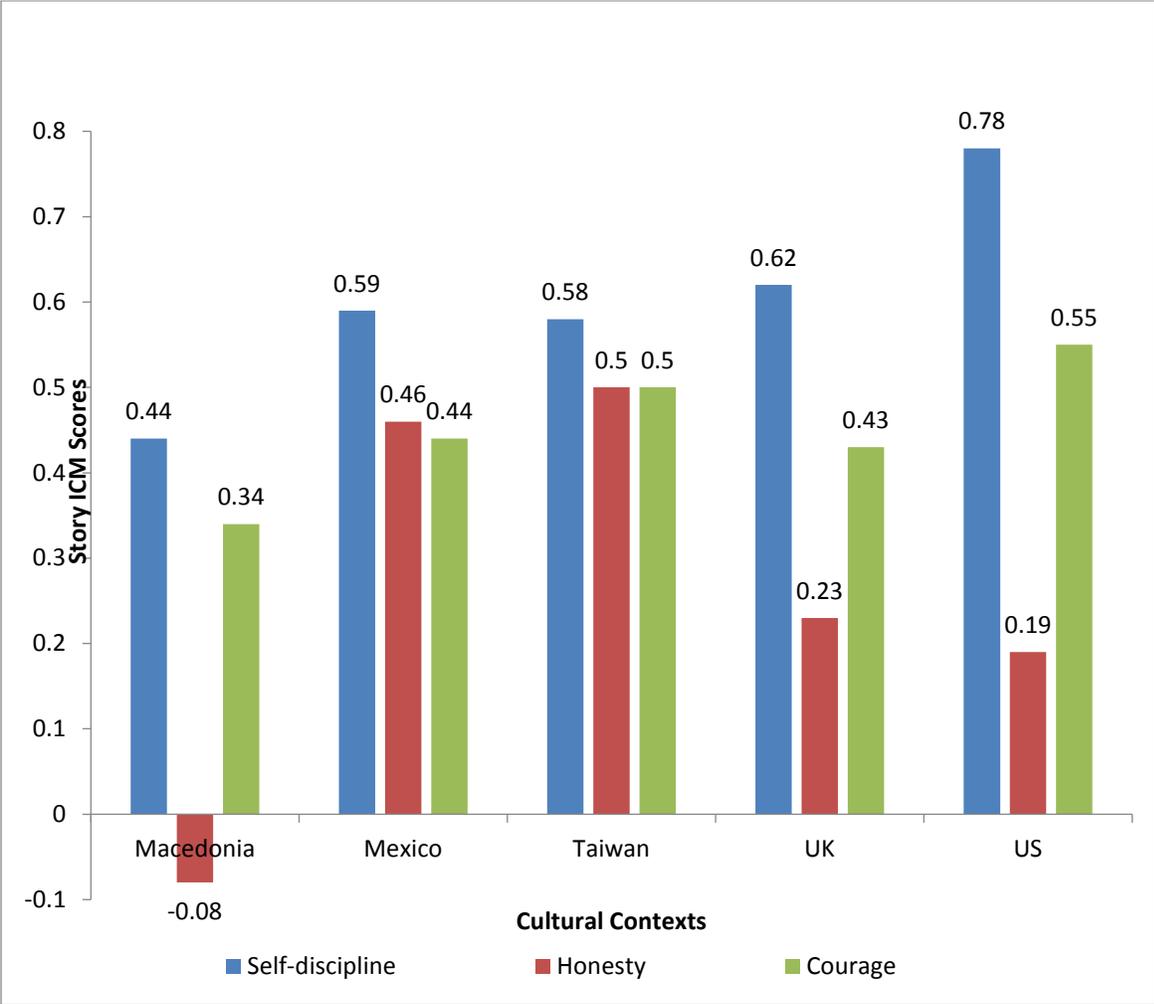


Figure3. Age differences by action choice and justification sub-scales.

