Check for updates

RESEARCH ARTICLE

REVISED Development and validation of the English version of the Moral Growth Mindset measure [version 3; peer review: 3 approved, 1 approved with reservations]

Hyemin Han [iD]1*, Kelsie J. Dawson [iD]1*, YeEun Rachel Choi[1], Youn-Jeng Choi [iD]2, Andrea L. Glenn[3]

[1]Educational Psychology Program, University of Alabama, Tuscaloosa, AL, USA
[2]Educational Research Program, University of Alabama, Tuscaloosa, AL, USA
[3]Center for the Prevention of Youth Behavior Problems, University of Alabama, Tuscaloosa, AL, USA

* Equal contributors

Abstract

Background: Moral Growth Mindset (MGM) is a belief about whether one can become a morally better person through efforts. Prior research showed that MGM is positively associated with promotion of moral motivation among adolescents and young adults. We developed and tested the English version of the MGM measure in this study with data collected from college student participants.
Methods: In Study 1, we tested the reliability and validity of the MGM measure with two-wave data (N = 212, Age mean = 24.18 years, SD = 7.82 years). In Study 2, we retested the construct validity of the MGM measure once again and its association with other moral and positive psychological indicators to test its convergent and discriminant validity (N = 275, Age mean = 22.02 years, SD = 6.34 years).
Results: We found that the MGM measure was reliable and valid from Study 1. In Study 2, the results indicated that the MGM was well correlated with other moral and positive psychological indicators as expected.
Conclusions: We developed and validated the English version of the MGM measure in the present study. The results from studies 1 and 2 supported the reliability and validity of the MGM measure. Given this, we found that the English version of the MGM measure can measure one's MGM as we intended.

Keywords

moral growth mindset, growth mindset, reliability, validity, moral development

Open Peer Review

Reviewer Status ✓ ? ✓ ✓

	Invited Reviewers			
	1	**2**	**3**	**4**
version 3 (revision) 21 Jul 2020				
version 2 (revision) 07 May 2020	✓ report		✓ report	✓ report
	↑			
version 1 14 Apr 2020	? report	? report		

1 **Michael T. Warren** [iD] , University of British Columbia, British Columbia, Canada
Western Washington University, Bellingham, USA

2 **Susan Mangan**, Claremont Graduate University, Claremont, USA
Thrive Center for Human Development, Pasadena, USA

3 **Elina Kuusisto**, Tampere University, Tampere, Finland

Kirsi Tirri, University of Helsinki, Helsinki, Finland

4   **Di You**, Alvernia University, Reading, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Hyemin Han (hyemin.han@ua.edu)

Author roles: Han H: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Dawson KJ**: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Writing – Original Draft Preparation, Writing – Review & Editing; **Choi YR**: Conceptualization, Data Curation, Investigation, Methodology, Writing – Review & Editing; **Choi YJ**: Methodology, Writing – Review & Editing; **Glenn AL**: Investigation, Writing – Review & Editing

Introduction

In the present study, we aimed to create and validate the English version of the Moral Growth Mindset (MGM) measure, which was originally developed in Korean. *Growth mindset* refers to the belief that it is possible to improve one's abilities and qualities, such as intelligence or personality[1]. These individuals believe that this can be done through effort and learning, which helps fosters motivation. Higher motivation for those with a growth mindset is encouraged through having attitudes such as viewing hardships as a chance to work harder rather than an indication of failure, and striving for success due to genuinely wanting to learn instead of being concerned with how others view them[2]. One study found that an intervention that taught students how to endorse a growth mindset reduced levels of aggression as well as depressive symptoms that resulted from being a victim of bullying[3]. This study suggested that growth mindset might be beneficial for promoting a sense of resilience when faced with social challenges or other difficulties.

MGM refers to growth mindset in the domain of morality. This mindset is related to one's belief that it is possible to become a morally better person and improve one's morals through efforts. A previous study showed that MGM was positively associated with increases in voluntary service engagement among adolescents and young adults[4]. The results suggested that among younger populations, MGM might increase participants' prosocial behavior due to the belief that it will make them morally better. Given this, MGM would be considered as a factor that contributes to moral development. In order to adequately examine how MGM contributes to moral development, however, it is necessary to have an appropriate measure. Additionally, if moral growth mindset motivates people to learn how to become more moral, as previous research suggests, then it is important for moral educators to have a tool to assess the malleability beliefs students have related to their morals. For example, if moral educators are able to identify that some students have a fixed mindset related to their morals, then an appropriate starting point may be to provide them with evidence that it is possible to improve moral character throughout one's life.

Although how MGM influences moral growth among people have been examined in several previous studies with a quantitative measure (e.g., 1,5,6), we found some points that could be improved in them, so we intended to develop and test our updated MGM measure for future studies. MGM was previously included as a three-item subscale in a general measure of growth mindset called the Theory Measures[5,6]. However, because it is important to include four or more items per factor to perform psychometric tests[7], the psychometrical qualities of the MGM subscale could not be sufficiently tested. For instance, the aforementioned previous studies examined the MGM as a subscale, so they could not sufficiently examine its internal structure and its association with diverse moral and positive psychological indicators. In a previous study[4], we developed and tested a Korean version of the MGM measure and evaluated the internal consistency and structure of the measure. However, the test-retest consistency and discriminant validity of the measure were not examined. Hence, in the present study, we created an English version of the MGM measure and tested its psychometric properties. In Study 1, we tested the internal and test-retest consistency and validity of the MGM measure and modified the measure to improve the model fit. In Study 2, we examined correlations between the MGM and other moral and positive psychological indicators associated with positive youth development to test the convergent and discriminant validity of the measure.

Study 1

In Study 1, we translated the MGM measure to English and tested its reliability and validity with two-wave data. We also modified the items to improve the model fit.

Methods

Translation of the MGM measure to English. Based on the Korean version of the MGM measure[4] and the Implicit Theory measure[1,8], we developed the English version of the MGM measure. Although the English version was created based on the Korean version, we did not do direct translation because of cultural differences in concepts and terms related to morals and characters (e.g., 9). Instead, the inventors (HH, KJD, and YJC) of the Korean MGM measure created its English version based on the structure of the Korean version and the wording in the Implicit Theory measure. In addition, the Implicit Theory measure was used due to the fact that it had six items and was based on Dweck's original measure of growth mindset for intelligence. As a result, the tested measure included six items as well (e.g., "No matter who you are, you can significantly improve your morals and character") and answers were anchored to a six-point Likert scale (i.e., strongly disagree (1), disagree (2), mostly disagree (3), mostly agree (4), agree (5), strongly agree (6)) (see *Extended data* for the full measure[10]).

Although Chiu, Hong, and Dweck[11] originally used more nuanced keywords such as "responsible and sincere" as well as "conscientiousness, uprightness, and honesty," we decided to use the more general terms, "morals and character." This was due to the concern that such nuanced terms in the original measure may be associated with specific moral foundations and biased towards certain groups of people. For example, conservatives have been found to score higher on measures of conscientiousness[12] whereas liberals have been found to rely primarily on the value of fairness, which is closely related to honesty, when dealing with moral issues (see research on Moral Foundation Theory; e.g., 13). Thus, we used "morals and characters" in order for participants to be able to define the terms based on

their own experiences and understanding. Finally, since Chiu *et al.* (1994)[11] used terms related to specific morals and characteristics in their original three-item subscale (e.g., "A person's moral character," "whether a person is responsible and sincere," "a person's moral traits"), we decided to use "morals and character" in order to stay consistent with the construct they were measuring. That is, rather than measuring participants' malleability beliefs about the overarching system of values they have, we wanted to measure malleability beliefs regarding individual morals, as did the original measure. Doing so may increase the chance for interventions since if people want to become a better person (improve their morality) they may need to believe that their values (morals) can be improved.

There might be another concern regarding use of the terms, "morals and characters," in our measure related to whether a simple belief about the possibility to improve ones' morals (or moral values) is directly relevant to their moral growth. For instance, one previous study about the relationship between one's endorsed personal values and behavioral outcomes reported that the correlation between values and behaviors, particularly other-reported behaviors, was moderate at the greatest[14]. Despite of this potential issue, however, we decided to use the terms because we tried to design the measure so that participants intuitively perceive and interpret items about "morals and characters" while considering their own beliefs about moral growth. In fact, researchers in implicit theories proposed that both incremental and entity theories are intuitive and "implicit" to people[15], so we intended to develop our items based on this point.

Our measure is in line with the original measure, the Implicit Theory measure, consisting of six items[1]. In fact, although all of the items were meant to measure whether or not participants endorse a growth mindset and are similar to each other, the wordings varied slightly to include core concepts of growth mindset such as being able to improve regardless of who you are (i.e., "no matter who you are"), the point in time (i.e., "always"), or the degree (i.e., "considerably"). In addition, because we were interested in whether MGM can be differentiated from growth mindset in general measured by the original growth mindset measure, we decided to use the same terms and format that were adopted in the original measure (e.g., "No matter who you are, you can change your intelligence a lot").

Participants. Study 1 was conducted during the 2018 fall semester. Participants were recruited from students enrolled in undergraduate educational psychology classes and they were provided with a course credit. Only students who were at least 18 years of age were eligible to complete the survey. The participants visited the subject pool system, checked the list of active research projects, and selected and signed up for our study. We decided to recruit at least 200 participants since $N = 200$ has been regarded as the recommended minimum sample size for confirmatory factor analysis (CFA)[16].

A total of 212 college students (89.15% females; Age mean = 24.18 years, SD = 7.82 years; 177 Caucasian, 34 African American, 1 Native American, 1 Asian, 1 Pacific Islander, 3 Latinx, 2 multi-ethnic) from the southern USA completed the

English MGM measure online via Qualtrics. They were re-invited to complete the same survey again one week later ($N = 207$ for Wave 2; 89.37% females; Age mean = 24.28 years, SD = 7.88 years).

Procedures. Participants who voluntarily signed up for study 1 received a link to the Qualtrics survey where they completed the MGM measure, followed by a demographics survey. When the participants signed up for the study, the subject pool manager provided us with their email addresses, and we sent the participants the survey link via email. We created our Qualtrics survey in a way that only the participants who answered all survey questions were able to complete the survey and receive a credit for their class. Thus, there was no missing data in the present study.

A consent form was sent out to the students alongside the MGM measure. This form was reviewed by the Institutional Review Board at the University of Alabama (IRB approval number: 18-04-1156), along with the approved studies, and was presented at the beginning of the Qualtrics form. Only students who read the form and agreed to participate in this study were presented with the survey forms.

Analysis. When examining test-retest reliability, we excluded participants who failed to complete the second survey within two weeks to control for the time gap between the two surveys, which left 168 cases for examining test-retest reliability (Mean time gap between Waves 1 and 2 = 7.78 days, SD = 1.66 days).

First, we examined consistency indices, i.e., Cronbach's α and test-retest consistency. Second, we performed CFA to examine the internal structure of the measure. We used robust weighted least squares (WLSMV) because it is more suitable for testing Likert-type items in a small sample[17]. During this process, we checked whether any item should be excluded from the measure to achieve a good model fit. If the measure was modified, we calculated all reliability and validity indices again. We used R (3.6.1) for statistical analyses. All data files and source codes are available as *Underlying data*[10].

Results

First, the measure demonstrated at least acceptable reliability (> .7; see Table 1) according to both Cronbach's alpha values and test-retest reliability. Second, we performed CFA – the original model with all six items did not show good model fit (see Table 1). Thus, we excluded items 1 and 2 while referring to Han *et al.* (2018), because in that study we showed relatively lower factor loadings in the six-item and five-item models respectively. In the supplementary table in *Extended data*, we presented factor loadings for the six-item and five-item models. In the six-item model, Item 1 showed the lowest standardized factor loading, identical to what was reported in Han *et al.* (2018)[4]. After excluding Item 1, Item 2 showed the lowest standardized loading in the five-item model, so we removed this item accordingly.

The CFA demonstrated that the four-item model was the best model given excellent model fit indices (chi-square test

p-value > .05, RMSEA and SRMR < .05, TLI and CFI > .95; see Table 2 for the best model[18]). In addition, as shown in Table 1, when we recalculated reliability indices , Cronbach α and test-retest *r*, after exclusion of the items, they all remained greater than .7.

In addition to the low factor loadings, we also decided to remove items 1 and 2 due to the fact that they may have been too vague. For example, item 1 stated "you can't really do much" and item 2 stated "you can't improve very much" whereas the other items used words such as "significantly improve," "always substantially improve," and "improve…considerably" that conveyed more specific magnitude. Using the less extreme terms in items 1 and 2 may have put the items at risk of inconsistency[19] since it would be easier for participants' opinions to shift regarding whether or not you can change "much." In addition, as another possibility, items 1 and 2 are more likely about entity beliefs, not malleability beliefs that constitute the basis of growth mindset. These items contain some words perhaps related to entity beliefs (e.g., "certain morals and characters...," "something about you…"), so they might not directly measure the core of the growth mindset construct and showed lower factor loadings compared to the other items.

Study 2
In Study 2, we tested the correlation between MGM and other moral and positive psychological indicators associated with positive youth development. In addition, we performed CFA for model confirmation. We aimed at testing the validity of the measure, construct, convergent, and divergent validity.

We selected several moral and positive psychological measures to test the convergent and divergent validity of the MGM measure. We employed the Implicit Theory Measure[1], which measures growth mindset in general, particularly intelligence growth mindset and constitutes the basis of the MGM measure, to test convergent and discriminant validity. For the selection of moral psychological measures, we referred to recent articles about psychological constructs that significantly predict prosocial and civic behavior[20]. They proposed moral judgment[21,22], moral emotion (empathy)[23], and moral identity[24] as fundamental constructs in moral functioning. We also employed the Propensity to Morally Disengage Scale to examine whether the MGM showed negative correlation with moral disengagement[25] since Han *et al.* (2018)[4] reported that MGM promotes moral engagement. In addition to the aforementioned moral psychological measures, we used the Claremont Purpose Scale as a way to examine

Table 1. MGM measure English reliability check and confirmatory factor analysis (CFA) results.

Model	Reliability		Classical CFA						
	Cronbach α	Test-retest *r*	χ2	*df*	*p*	CFI	TLI	RMSEA	SRMR
Study 1									
6-item	.86	.76	60.08	9	.000	.84	.73	.16	.09
5-item (without item 1)	.86	.74	26.73	5	.000	.92	.83	.14	.07
4-item (without items 1 and 2)	.85	.70	1.79	2	.41	1.00	1.00	.00	.01
Study 2									
4-item (without items 1 and 2)	.77	-	1.60	2	.45	1.00	1.00	.00	.01

Table 2. Factor loadings from CFA in both studies.

Item	Study 1		Study 2	
	Unstandardized	Standardized	Unstandardized	Standardized
No matter who you are, you can significantly improve your morals and character.	.72	.69	.77	.66
To be honest, you can't really improve your morals and character.	-.73	-.73	-.46	-.39
You can always substantially improve your morals and character.	.75	.75	.89	.81
You can improve your basic morals and character considerably.	.86	.89	.93	.94

one's positive development in terms of flourishing[26], given that purpose has been regarded as a possible moral virtue for eudemonic wellbeing[27].

In general, according to the previous studies that examined the relationship between growth mindset, positive psychological indicators, and antisocial tendency (e.g., [28]–[30]), we hypothesized that the sizes of correlation coefficients between MGM and other indicators, except the intelligence growth mindset, would be between .10 (small) and .30 (medium). We discussed further details regarding the hypothesized effect size of each measure in the following sections.

Methods

Participants. As per Study 1, participants were recruited from the educational psychology and psychology subject pools during the 2019 spring semester, with similar age and class enrollment restrictions employed in Study 1, Participants in educational psychology classes visited the subject pool system, checked the list of active research projects, and selected and signed up for our study. Participants in psychology classes who intended to sign up for our study visited the SONA system, reviewed the list of active studies, and then selected and signed up for the present study.

In total, 275 college students (81.45% females; Age mean = 22.02 years, SD = 6.34 years; 223 Caucasian, 39 African American, 2 Native American, 1 Asian, 1 Pacific Islander, 5 Latinx, 4 multi-ethnic) in the Southern United States of America were recruited. The consent procedure was identical to that in Study 1 (The University of Alabama IRB approval numbers: 18-10-1633, 18-12-1842).

Procedures. When participants signed up for the present study, the procedure for educational psychology students was identical to that of study 1. In the case of psychology students, they were automatically provided with a link to a Qualtrics survey via the SONA system. Participants were presented with the MGM measure and other moral and positive psychological measures, all of which were presented in a randomized order, followed by a demographics survey. Similar to Study 1, only the participants who answered all questions were able to complete the survey and receive a credit, so there was no missing data in the present study. For sample size estimation, similar to Study 1, we followed the guidelines for CFA[16], so we determined that at least 200 participants were required.

Measures. MGM measure. We used the four-item MGM measure used in Study 1.

Implicit Theory Measure. The Implicit Theory Measure was designed to measure one's mindset regarding whether it is possible to change and improve one's intelligence and abilities in general[1]. The measure consists of six items and responses are anchored to a six-point Likert scale. The structure of this measure has been tested in previous studies (e.g., 1, 8). Given that the Implicit Theory Measure measures one's general growth mindset, we expected that it would be positively

correlated with MGM. However, because the construct measured by the Implicit Theory Measure is not domain specific, we also expected that the MGM would not completely overlap with this construct (discriminant validity). Given these, the effect size of the correlation coefficient would be medium to large ($r = +.3 - +.5$).

Behavioral Defining Issues Test (bDIT). The bDIT was developed to assess development of one's moral judgment[21,22]. Choi *et al.* (2019)[21] tested its measurement structure and psychometrical qualities and found that it did not favor any gender and it showed acceptable reliability as well as concurrent validity with the DIT-1 measure. In general, the bDIT assesses whether one can make moral judgments based on the post-conventional schema instead of focusing on social norms or one's personal interests. It consists of three moral dilemmas and 24 questions that ask what the most important moral philosophical criterion is when solving the moral dilemmas. We used a percentile score that quantified the likelihood of utilizing the post-conventional schema. Because the bDIT measures one's moral judgment development, we expected that MGM would be positively associated with the bDIT score.

Unlike other self-report measures, the bDIT is a behavioral measure evaluating one's developmental level of moral judgment with behavioral responses. Previous research has shown that participants could not increase their score even if they were asked to fake higher moral judgment with the DIT[31]. Thus, the bDIT is less susceptible to social desirability bias and can measure one's actual moral functioning instead of self-reported qualities. Given that this is a psychological test to assess one's moral functioning and not a self-report measure, we expected that the bDIT score would be weakly correlated with MGM ($r \sim +.1$).

Interpersonal Reactivity Index (IRI). The IRI was used to measure empathic traits, i.e., empathic concern (EC), personal distress (PD), perspective taking (PT), and fantasy scale (FS) (Davis, 1983) with 28 items. The internal structure of the measure based on the four-factor model was validated in previous studies with factor analysis (see Chrysikou & Thompson, 2016). According to Decety and Cowell's (2014) discussion regarding the relationship between different subcomponents in the IRI and moral functioning, we hypothesized that only EC and PT, not PD and FS, would be positively correlated with MGM[32]. Given the IRI is a self-report measure, we expected a relatively larger (small to medium) effect size of correlation, $r = +1 - +3$, compared with the bDIT.

Moral Identity Scale (MIS). The MIS measures moral identity in terms of whether moral values are regarded as central to one's self-identity[24]. Five items measure the moral internalization subscale and six items measure the moral symbolization subscale. Aquino and Reed (2002)[24] also performed CFA to validate its internal structure. Given that previous research showed that moral internalization is more fundamental in predicting one's internal moral belief and motivation[24], we hypothesized that only moral internalization would be significantly associated with MGM. The hypothesized effect size of the correlation

would be similar to that of the correlation between MGM, EC, and PT ($r = +.1 - +.3$).

Propensity to Morally Disengage Scale. The moral disengagement scale measures one's propensity to disengage from moral behavior within morally problematic situations[25]. It measures moral disengagement propensities for eight mechanisms (i.e., moral justification, euphemistic labeling, advantageous comparison, displacement of responsibility, diffusion of responsibility, distortion of consequences, dehumanization, attribution of blame) with eight items (one item per mechanism). We used a composite score of the eight items. The internal structure of the scale was tested with CFA by Moore *et al.* (2012)[25]. As Bandura (2002) proposed[33], moral disengagement is negatively associated with motivation for moral engagement. Thus, we expected moral disengagement would be negatively associated with MGM while the effect size of the correlation would be similar to the cases of the IRI and MIS (small to medium; $r = -.1 - .3$).

Claremont Purpose Scale (CPS). This 12-item scale quantitatively measures purpose among adolescents using three subscales: meaningfulness, goal orientation, and beyond-the-self dimension[26]. CPS scores were positively associated with various moral and positive psychological indicators (e.g., purpose in life, satisfaction with life, empathic concern, wisdom) in prior research[26]. We used both the total CPS and subscale scores given that Bronk et al. (2018)[26] validated it with hierarchical CFA. Given

previous studies that examine the association between morality, meaning[34], and purpose[33,35], similar to the cases of the IRI and MIS, we hypothesized a small to medium effect size of the correlation between MGM and CPS ($r = +1 - +3$).

Analysis. First, we performed CFA with the MGM data again to test the internal structure of the MGM measure (construct validity). Second, we conducted correlation analyses to examine how MGM was associated with other moral and positive psychological indicators (convergent validity). Third, we tested whether or not the MGM measure examines a construct independent from intelligence growth mindset (discriminant validity) using the Fornell-Larcker criterion[36].

We also used R in Study 2. All data files and source codes are available as *Underlying data*[10].

Results
The results of the reliability check showed that the MGM measure as well as all other measures possessed at least acceptable reliability (> .7; see Table 3). Moreover, CFA supported good internal structure of the MGM measure (see Table 1 and Table 2). However, it should be acknowledged that Item 4 showed a lower factor loading in Study 2 compared with Study 1, although the overall model fit indices were excellent. It would be an issue since researchers have regarded .40 as the threshold for a good factor loading[37]. Although the issue could not be resolved completely with the current dataset, one larger dataset ($N = 701$

Table 3. Descriptive statistics, Cronbach's α, and correlation test results.

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	4.77	.86	.77													
2	4.36	1.03	.37***	.90												
3	51.23	21.14	.11†	.15*	.78											
4	3.84	.68	.25***	.22***	.18**	.78										
5	2.76	.66	-.06	.04	-.07	.02	.70									
6	3.61	.62	.26***	.14*	.23***	.58***	-.03	.70								
7	3.47	.78	.16*	.22***	.19**	.33***	.22***	.18**	.77							
8	4.44	.72	.34***	.25***	.20**	.53***	-.09	.38***	.23***	.80						
9	3.31	.85	.04	.08	-.15*	.09	.14*	.13*	.08	.10	.86					
10	2.40	1.09	-.24***	-.28***	-.15*	-.34***	.11†	-.22***	-.14*	-.37***	-.07	.88				
11	3.83	.63	.16**	.16**	.01	.27***	-.16**	.25***	.15*	.24***	.23***	-.13*	.89			
12	3.52	.94	.06	.07	-.13*	.05	-.23***	.09	.05	.06	.24***	-.02	.82***	.90		
13	4.00	.70	.22***	.17**	.08	.22***	-.10	.17**	.16**	.21***	.07	-.12†	.79***	.49***	.86	
14	3.96	.77	.12*	.16**	.10	.40***	-.02	.34***	.16**	.34***	.21***	-.19**	.74***	.35***	.45***	.86

Note. M: mean. *SD*: standard deviation. *r*: Pearson correlation coefficient. Cronbach αs are also reported (on the diagonal).

† *p* < .10

* *p* < .05, ** *p* < .01, *** *p* < .001. 1: MGM, 2: intelligence growth mindset, 3: bDIT, 4: IRI EC, 5: IRI PD, 6: IRI PT, 7: IRI FS, 8: moral internalization, 9: moral symbolization, 10: moral disengagement, 11: CPS all, 12: CPS meaningfulness, 13: CPS goal orientation, 14: CPS beyond-the-self dimension.

as of May 2020) that is currently being collected for the next research project was analyzed as a possible way to address the issue. When we conducted preliminary CFA with the new dataset, all four factor loadings were greater than .75 while the CFA model showed good model fit, RMSEA = .08, SRMR = .01, TLI = .98, CFI = .99. Given this, the small factor loading of Item 4 reported in Study 2, -.39, could be addressed in the long term with additional data collection and analysis.

Correlation analysis demonstrated a positive association between MGM, intelligence growth mindset, and other moral psychological indicators such as empathic concern, perspective taking, moral internalization, and purpose. Indicators relatively less relevant to morality, such as personal distress, symbolization, and meaningfulness, did not show a significant correlation (see Table 3). The effect size of the correlation coefficient between MGM and bDIT was small as predicted, but the correlation was non-significant ($p = .08$). MGM was not significantly correlated with PD and CPS meaning. The correlation between MGM and moral disengagement was significantly negative. We found that the correlation coefficient between MGM and intelligence growth mindset ($r = .37$) was smaller than the square root of the average variance extracted (AVE=.84), which indicates MGM showed discriminant validity from intelligence growth mindset.

Discussion

We developed and tested the English version of the MGM measure in this study with data collected from emerging adult participants. In Study 1, we found that the four-item MGM measure possessed good consistency and internal structure. In fact, the previous studies that developed and tested measurements for diverse types of domain-specific growth mindset have shown that the measurements possessed good reliability and validity as well (e.g., 38,39). Consistent with these previous studies, we were able to show that MGM can also be appropriately measured by a self-report measure, the English version of the MGM measure, as we intended.

In Study 2, we found that MGM was positively associated with moral and positive psychological indicators as hypothesized. Two exceptions were the significant associations between MGM and FS and the non-significant association between MGM and CPS meaning. First, FS is intended to quantify one's tendency to expand their empathy toward imaginary beings, so the significant association with MGM indicates a tendency to broaden one's empathy. Second, CPS meaning is about personal meaning, which does not necessarily always mean moral[40], so it makes sense that it would not be significantly associated with MGM. This result would suggest that the MGM measures a construct that is specifically about moral development in addition to positive youth development in general.

In the case of the bDIT, the effect size was within the hypothesized range, but the correlation was non-significant ($p = .08$) perhaps due to the small sample size. As previously mentioned,

this could also be due to the fact that the bDIT is a behavioral measure rather than a self-report measure like the MGM measure. Since the bDIT is less susceptible to social desirability bias, it may be necessary to further explore the possibility of bias in participants' responses for the MGM measure in future studies.

In addition, moral disengagement was negatively correlated with MGM. Since moral disengagement allows people to dismiss negative feelings, they may have about behaving immorally using the eight mechanisms previously mentioned, this increases the likelihood of continuing to behave immorally. In this way, moral disengagement and MGM have somewhat reverse trajectories. As hypothesized, this suggests that MGM may promote engaging in moral behavior. In addition, since moral internalization, which has been shown to inhibit moral disengagement[41], was also positively correlated with MGM, it makes sense that our measure was negatively correlated with moral disengagement. If somebody has a strong sense of their morals and these values are internalized, this may help them to stay engaged with their standards and furthermore, be motivated to continue to be morally better.

Finally, we found good discriminant validity between the MGM measure and the intelligence growth mindset measure as a measure for growth mindset in general. This indicates that although the intelligence growth mindset measure and the MGM measure are measuring growth mindset related to different domains, they are measuring distinctly different constructs related to malleability beliefs (i.e., intelligence and morals, respectively). Given this, our MGM measure significantly contributes to growth mindset research by introducing a reliable and valid measure for growth mindset related to morals.

The results from our correlation analysis are consistent with findings in previous studies that have examined the positive relationship between growth mindset and successful social adjustment and positive youth development in general[2,28,42].

This English version of the MGM measure has the potential to significantly contribute to research in moral development and education. For instance, researchers and educators who are interested in how MGM is associated with moral development may use the MGM measure in their studies. In addition, given that we created the English version of the MGM measure, scholars who are using languages other than Korean or English will be able to translate the measure into their languages. By doing so, it would be possible to accumulate large-scale datasets for testing the measure in diverse backgrounds and contexts, and to examine the roles of MGM in moral development in the long term.

However, there are limitations in this study that warrant future studies. First, we collected data only from undergraduate students and male students were underrepresented in both studies; such issues may limit the generalizability of our findings. Second, although we used straightforward terms (e.g., morals

and characters), we could not test whether the measure was actually unbiased according to one's political orientation of endorsed moral foundations. To address this issue, measurement invariance test would be a way to examine whether the MGM measure, which allows participants to interpret "morals" and "characters" by themselves, measures the same construct across different groups who may use different underlying folk conceptions of morals and characters. Third, although participants spent about 33.98 minutes (median) to complete Study 2 we did not include any attention check items. Fourth, we did not employ Chiu et al.'s (1997)[5] original measure, which could be informative while conducting the convergent validity check, although our measure was based on Dweck's (2000)[1] updated six-item intelligence growth mindset measure. Fifth, the items used in the MGM measure could be revised particularly when being administered among younger populations. We decided to use the current wordings to maintain consistency with the Korean version of the MGM measure and the Implicit Theory Measure, which constituted the basis of our measure. However, to make the measure more applicable to younger populations, some complex words (e.g., "substantially," "considerably") could be replaced with simpler words (e.g., "a lot"). Sixth, since several items in the measure might seem to be similar, the words could be revised in future studies, particularly those focusing on children or young adolescents. Although we decided to use the items overlapping with each other to make our measure consistent with previous growth mindset measures, further studies are needed to examine which form of the measure would be more appropriate. Seventh, the MGM has been tested only in the United States (mainly among Caucasians) and Korea, so it needs to be tested in more diverse countries and ethnicities to examine its cross-cultural validity.

Data availability

Open Science Framework: Moral Growth Mindset is Associated with Change in Voluntary Service Engagement, https://doi.org/10.17605/OSF.IO/VMJUA[10].

Underlying data

Folder "English version MGM" contains the underlying data and source code files that support the findings of this study:

- DISC.csv
- DISC_SONA.csv
- post.csv
- pre.csv

Extended data

Folder "English version MGM" contains the following extended data files:

- README
- EJDP.R
- Supplementary Materials.docx- MGM measure in English and information about additional moral and positive psychological measures used in Study 2
- Supplementary Table.xlsx-Supplementary table reporting factor loadings from 6-item and 5-item models (contained in folder 'English version MGM', Supplementary table.xlsx)

References

1. Dweck CS: **Self-Theories: Their Role in Motivation, Personality, and Development.** Philadephia, PA: Psychology Press. 2000.
   Reference Source

2. Yeager DS, Dweck CS: **Mindsets that promote resilience: When students believe that personal characteristics can be developed.** *Educ Psychol.* 2012; **47**(4): 302–14.
   Publisher Full Text

3. Yeager DS, Trzesniewski KH, Dweck CS: **An implicit theories of personality intervention reduces adolescent aggression in response to victimization and exclusion.** *Child Dev.* 2013; **84**(3): 970–88.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. Han H, Choi YJ, Dawson KJ, *et al.*: **Moral growth mindset is associated with change in voluntary service engagement.** *PLoS One.* Lamm C, editor. 2018; **13**(8): e0202327.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

5. Chiu C, Dweck CS, Tong JY, *et al.*: **Implicit theories and conceptions of morality.** *J Pers Soc Psychol.* 1997; **73**(5): 923–40.
   Publisher Full Text

6. Dweck CS, Chiu C, Hong Y: **Implicit theories and their role in judgments and reactions: A word from two perspectives.** *Psychol Inq.* 1995; **6**(4): 267–85.
   Publisher Full Text

7. Netemeyer R, Bearden W, Sharma S: **Scaling Procedures.** Thousand Oaks, CA: SAGE Publications, Inc. 2003.
   Publisher Full Text

8. Burnette JL: **Implicit theories of body weight: entity beliefs can weigh you down.** *Pers Soc Psychol Bull.* 2010; **36**(3): 410–22.
   PubMed Abstract | **Publisher Full Text**

9. Han H, Glover GH, Jeong C: **Cultural influences on the neural correlate of moral decision making processes.** *Behav Brain Res.* 2014; **259**: 215–28.
   PubMed Abstract | **Publisher Full Text**

10. Han H, Choi Y, Dawson KJ, *et al.*: **Moral growth mindset is associated with change in voluntary service engagement.** 2020.
    http://www.doi.org/10.17605/OSF.IO/VMJUA

11. Chiu C, Hong Y, Dweck CS: **Toward an integrative model of personality and intelligence: A general framework and some preliminary steps.** In: Sternberg RJ, Ruzgis P, editors. *Personality and Intelligence.* Cambridge, UK: Cambridge University Press. 1994; 104–34.
    Reference Source

12. Schlenker BR, Chambers JR, Le BM: **Conservatives are happier than liberals, but why? Political ideology, personality, and life satisfaction.** *J Res Pers.* 2012; **46**(2): 127–46.
    Publisher Full Text

13. Graham J, Haidt J, Nosek BA: **Liberals and conservatives rely on different sets of moral foundations.** *J Pers Soc Psychol.* 2009; **96**(5): 1029–46.
    PubMed Abstract | **Publisher Full Text**

14. Plaks JE, Grant H, Dweck C: **Violations of implicit theories and the sense of prediction and control: Implications for motivated person perception.** *J Pers Soc Psychol.* 2005; **88**(2): 245–262.
    Publisher Full Text

15. Bardi A, Schwartz SH: **Values and behavior: Strength and structure of relations.** *Pers Soc Psychol Bull.* 2003; **29**(10): 1207–1220.
    PubMed Abstract | **Publisher Full Text**

16. Gagne P, Hancock GR: **Measurement Model Quality, Sample Size, and Solution Propriety in Confirmatory Factor Models.** *Multivariate Behav Res.* 2006; **41**(1):

65–83.
PubMed Abstract | **Publisher Full Text**

17. Flora DB, Curran PJ: **An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data.** *Psychol Methods.* 2004; **9**(4): 466–91.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Hu L, Bentler PM: **Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.** *Struct Equ Model A Multidiscip J.* 1999; **6**(1): 1–55.
Publisher Full Text

19. Hong Y, Chiu C, Dweck CS, *et al.*: **Implicit theories, attributions, and coping: A meaning system approach.** *J Pers Soc Psychol.* 1999; **77**(3): 588–99.
Publisher Full Text

20. Darnell C, Gulliford L, Kristjánsson K, *et al.*: **Phronesis and the Knowledge-Action Gap in Moral Psychology and Moral Education: A New Synthesis?** *Hum Dev.* 2019; **62**(3): 101–129.
Publisher Full Text

21. Choi YJ, Han H, Dawson KJ, *et al.*: **Measuring moral reasoning using moral dilemmas: evaluating reliability, validity, and differential item functioning of the behavioural defining issues test (bDIT).** *Eur J Dev Psychol.* 2019; **16**(5): 622–631.
Publisher Full Text

22. Han H, Dawson KJ, Thoma SJ, *et al.*: **Developmental level of moral judgment influences behavioral patterns during moral decision-making.** *J Exp Educ.* 2019.
Publisher Full Text

23. Davis MH: **Measuring individual differences in empathy: Evidence for a multidimensional approach.** *J Pers Soc Psychol.* 1983; **44**(1): 113–26.
Publisher Full Text

24. Aquino K, Reed A 2nd: **The self-importance of moral identity.** *J Pers Soc Psychol.* 2002; **83**(6): 1423–40.
PubMed Abstract | **Publisher Full Text**

25. Moore C, Detert JR, Klebe Trevino L, *et al.*: **Why employees do bad things: Moral disengagement and unethical organizational behavior.** *Pers Psychol.* 2012; **65**(1): 1–48.
Publisher Full Text

26. Bronk KC, Riches BR, Mangan SA: **Claremont Purpose Scale: A Measure that Assesses the Three Dimensions of Purpose among Adolescents.** *Res Hum Dev.* 2018; **15**(2): 101–17.
Publisher Full Text

27. Han H: **Purpose as a moral virtue for flourishing.** *J Moral Educ.* 2015; **44**(3): 291–309.
Publisher Full Text

28. Yeager DS, Trzesniewski KH, Tirri K, *et al.*: **Adolescents' implicit theories predict desire for vengeance after peer conflicts: correlational and experimental evidence.** *Dev Psychol.* 2011; **47**(4): 1090–107.
PubMed Abstract | **Publisher Full Text**

29. Diseth Å, Meland E, Breidablik HJ: **Self-beliefs among students: Grade level and gender differences in self-esteem, self-efficacy and implicit theories of intelligence.** *Learn Individ Differ.* 2014; **35**: 1–8.
Publisher Full Text

30. Tamir M, John OP, Srivastava S, *et al.*: **Implicit theories of emotion: affective and social outcomes across a major life transition.** *J Pers Soc Psychol.* 2007; **92**(4): 731–44.
PubMed Abstract | **Publisher Full Text**

31. Napier JD: **Effects of Knowledge of Cognitive-Moral Development and Request to Fake on Defining Issues Test P-Scores.** *J Psychol.* 1979; **101**(1): 45–52.
Publisher Full Text

32. Decety J, Cowell JM: **The complex relation between morality and empathy.** *Trends Cogn Sci.* 2014; **18**(7): 337–9.
PubMed Abstract | **Publisher Full Text**

33. Bandura A: **Selective moral disengagement in the exercise of moral agency.** *J Moral Educ.* 2002; **31**(2): 101–119.
Publisher Full Text

34. Han H, Liauw I, Kuntz AF: **Moral Identity Predicts the Development of Presence of Meaning During Emerging Adulthood.** *Emerg Adulthood.* 2019; **7**(3): 230–237.
Publisher Full Text

35. Han H, Ballard PJ, Choi YJ: **Links between moral identity and political purpose during emerging adulthood.** *J Moral Educ.* 2019.
Publisher Full Text

36. Henseler J, Ringle CM, Sarstedt M: **A new criterion for assessing discriminant validity in variance-based structural equation modeling.** *J Acad Mark Sci.* 2015; **43**(1): 115–35.
Publisher Full Text

37. Hair JF Jr, Black WC, Babin BJ, *et al.*: **Multivariate data analysis (7th edition).** Harlow, UK: Pearson Education Limited. 2010.
Reference Source

38. Lüftenegger M, Kollmayer M, Bergsmann E, *et al.*: **Mathematically gifted students and high achievement: the role of motivation and classroom structure.** *High Abil Stud.* 2015; **26**(2): 227–43.
Publisher Full Text

39. Pomerantz EM, Saxon JL: **Conceptions of ability as stable and self-evaluative processes: a longitudinal examination.** *Child Dev.* 2001; **72**(1): 152–73.
PubMed Abstract | **Publisher Full Text**

40. Bronk KC: **Purpose in life: A critical component of optimal youth development.** Dordrecht, The Netherlands: Springer. 2013.
Reference Source

41. Hardy SA, Bean DS, Olsen JA: **Moral Identity and Adolescent Prosocial and Antisocial Behaviors: Interactions with Moral Disengagement and Self-regulation.** *J Youth Adolesc.* 2015; **44**(8): 1542–54.
PubMed Abstract | **Publisher Full Text**

42. Yeager DS, Miu AS, Powers J, *et al.*: **Implicit theories of personality and attributions of hostile intent: a meta-analysis, an experiment, and a longitudinal intervention.** *Child Dev.* 2013; **84**(5): 1651–67.
PubMed Abstract | **Publisher Full Text**

Open Peer Review

Current Peer Review Status: ✔ ? ✔ ✔

Version 2

Reviewer Report 30 June 2020

✔ **Di You**

Psychology and Counseling Department, Alvernia University, Reading, PA, USA

The authors carefully designed two studies to test the reliability and validity of the English version of MGM, moral growth mindset. The literature review was relevant and coherent. The method section was well articulated. The results were clearly presented. The discussion was well-organized.
One recommendation for future study: given for both studies, majorities of the participants are Caucasian. A future study may want to investigate whether the MGM is consistent across various ethnicities.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: moral development; ethics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 16 Jul 2020

Hyemin Han, University of Alabama, Tuscaloosa, USA

Dear Dr. You,

Thank you very much for your positive comment and suggestion on our paper. Here is our response to your suggestion:

Suggestion: One recommendation for future study: given for both studies, majorities of the participants are Caucasian. A future study may want to investigate whether the MGM is consistent across various ethnicities.

Response: We appreciate your thought about the potential cross-cultural validity of our measure. We discussed the point in the discussion section:

Seventh, the MGM has been tested only in the United States (mainly among Caucasians) and Korea, so it needs to be tested in more diverse countries and ethnicities to examine its cross-cultural validity.

We appreciate your time and consideration once again.

Best,
Hyemin

Competing Interests: NA

Reviewer Report 16 June 2020

https://doi.org/10.5256/f1000research.26396.r64142

Elina Kuusisto
Tampere University, Tampere, Finland
Kirsi Tirri
Department of Education, University of Helsinki, Helsinki, Finland

The article is well written and constructed. The statistical analyses are sound and solid. Based on the results the instrument is reliable and valid.

1. However, we question what is the added value of the MGM instrument? In Dweck's (2000) previous work implicit beliefs about morality are investigated with her original instrument.

2. What is the relationship between morality and "morals and character"? What does morals mean here? What does character mean here? Why are these concepts combined?

3. The basic rule of items is that they should measure only one issue per one item.

4. In the limitations, author should reflect how to solve the issues mentioned above in the future studies.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: school pedagogy, moral education, gifted education, talent development

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 16 Jul 2020

Hyemin Han, University of Alabama, Tuscaloosa, USA

Dear Drs. Kuusisto and Tirri,

Thank you very much for your constructive comments and suggestions. They helped a lot while we were revising our manuscript. Here are our responses to your comments and suggestions:

1. However, we question what is the added value of the MGM instrument? In Dweck's (2000) previous work implicit beliefs about morality are investigated with her original instrument.

Response: We appreciate your comment about the contribution of our work. In the introduction section in the revised manuscript, we elaborated the point:

Although how MGM influences moral growth among people have been examined in several previous studies with a quantitative measure (e.g., 1, 5, 6), we found some points that could be improved in them, so we intended to develop and test our updated MGM measure for future studies. MGM was previously included as a three-item subscale in a general measure of growth mindset called the Theory Measures 5, 6 . However, because it is important to include four or more items per factor to perform psychometric tests 7 , the psychometrical qualities of the MGM subscale could not be sufficiently tested. For instance, the aforementioned previous studies examined the MGM as a subscale, so they could not sufficiently examine its internal structure and its association with diverse moral and positive psychological indicators.

2. What is the relationship between morality and "morals and character"? What does morals mean here? What does character mean here? Why are these concepts combined?

Response: Thanks a lot for your comment about the terms that we used. We agree with you that we used somehow general terms so they might not be ideal to examine specific concepts. However, we intended to examine one's MGM regardless of their ideas about morality and character, we decided to use those more general terms instead of specific terms. These general terms might be more suitable to examine people's folk and "implicit" concepts regarding their potential moral growth. In the revised manuscript, we discussed the point while responding to Reviewer 1's related comment:

Although Chiu, Hong, and Dweck 11 originally used more nuanced keywords such as "responsible and sincere" as well as "conscientiousness, uprightness, and honesty," we decided to use the more general terms, "morals and character." This was due to the concern that such nuanced terms in the original measure may be associated with specific moral foundations and biased towards certain groups of people. For example, conservatives have been found to score higher on measures of conscientiousness 12 whereas liberals have been found to rely primarily on the value of fairness, which is closely related to honesty, when dealing with moral issues (see research on Moral Foundation Theory; e.g., 13). Thus, we used "morals and characters" in order for participants to be able to define the terms based on their own experiences and understanding. Finally, since Chiu et al. (1994) 11 used terms related to specific morals and characteristics in their original three-item subscale (e.g., "A person's moral character," "whether a person is responsible and sincere," "a person's moral traits"), we decided to use "morals and character" in order to stay consistent with the construct they were measuring. That is, rather than measuring participants' malleability beliefs about the overarching system of values they have, we wanted to measure malleability beliefs regarding individual morals, as did the original measure. Doing so may increase the chance for interventions since if people want to become a better person (improve their morality) they may need to believe that their values (morals) can be improved.
There might be another concern regarding use of the terms, "morals and characters," in our measure related to whether a simple belief about the possibility to improve ones' morals (or moral values) is directly relevant to their moral growth. For instance, one previous study about the relationship between one's endorsed personal values and behavioral outcomes reported that the correlation between values and behaviors, particularly other-reported behaviors, was moderate at the greatest 40. Despite of this potential issue, however, we decided to use the terms because we tried to design the measure so that participants intuitively perceive and interpret items about "morals and characters" while considering their own beliefs about moral growth. In fact, researchers in implicit theories proposed that both incremental and entity theories are intuitive and

"implicit" to people 41, so we intended to develop our items based on this point.

At the same time, as we agree your point in general, we would like to mention that we briefly discussed the point in a part of the discussion section regarding limitations:

Second, although we used straightforward terms (e.g., morals and characters), we could not test whether the measure was actually unbiased according to one's political orientation of endorsed moral foundations. To address this issue, measurement invariance test would be a way to examine whether the MGM measure, which allows participants to interpret "morals" and "characters" by themselves, measures the same construct across different groups who may use different underlying folk conceptions of morals and characters.

We plan to examine how different people differently perceive the terms and concepts in our follow-up studies.

3. The basic rule of items is that they should measure only one issue per one item.
In the limitations, author should reflect how to solve the issues mentioned above in the future studies.

Response: We appreciate your comment regarding the items. In the limitation section, we discussed the point:

Sixth, since several items in the measure might seem to be similar, the words could be revised in future studies, particularly those focusing on children or young adolescents. Although we decided to use the items overlapping with each other to make our measure consistent with previous growth mindset measures, further studies are needed to examine which form of the measure would be more appropriate.

Again, we sincerely appreciate your time and consideration.

Best,
Hyemin

Competing Interests: N/A

Reviewer Report 11 May 2020

https://doi.org/10.5256/f1000research.26396.r63220

Michael T. Warren   iD

1 Human Early Learning Partnership, University of British Columbia, British Columbia, Canada

2 Psychology Department, Western Washington University, Bellingham, WA, USA

Dear Dr. Han,

Thank you for the opportunity to review the revision of your Moral Growth Mindset article. I appreciate your thorough attention to detail, and particularly the improvements that were made in fully reporting the CFA results and providing rationales for the selection of constructs used to test convergent and discriminant validity in Study 2. Accordingly, I am approving the article, as only minor changes are further warranted. Thank you for your valuable contribution to moral science.

I do remain skeptical about the word choice "improve your morals" found in the scale's items. My concern is that the phrase introduces unnecessary ambiguity. Are participants meant to interpret it as "improve yourself on specific moral traits" (based on your response I believe this is the intention), or are they meant to interpret it as "improve your moral values." To me, the former is directly relevant to becoming a better person, whereas the latter has very little to do with it-- we know, for example, that moral values and behaviour often show only small-to-moderate associations (e.g., Bardi & Schwartz, 2003[1]). Becoming a better person entails not only holding moral values but embodying the cognitive, affective, motivational, and behavioral expressions of one's moral values. That said, it's entirely possible that I'm overthinking this and that typical survey respondents will intuitively grasp the meaning you intend. I'll leave it to you to decide whether this issue warrants further discussion.

Beyond that, I would ask you to consider the following additional minor suggestions:
- Consider providing the MGM scale anchors either in the "Translation of the MGM measure to English" section (p. 3) or in a note beneath Table 2. I think requiring people to find the scale anchors in the "Supplementary Materials.docx" file may prove inconvenient; some future researchers may take the items from Table 2 and make up their own anchors.

- In discussing Item 4's smaller factor loading in Study 2, I don't think it's fair to say it had a "slightly lower factor loading in Study 2 compared with Study 1." A difference of -.73 vs. -.39 in standardized loadings is quite substantial. Additionally, I wonder if readers will know which item is #4, since you're referring to the second item in Table 2. This issue also pertains to other items mentioned by number throughout the document.

- Question: Is the Implicit Theory Measure in fact domain-general (as stated on p. 5) or does it measure intelligence growth mindset?

- I'm not sure what the following sentence means in the context of the model fit paragraph on p. 5: "As shown in Table 1, when we recalculated indices after exclusion of the items, they all remained greater than .7." I assume you're referring to alpha, but the context suggests you're saying the model fit indices were greater than .7, which is confusing. (It would also be confusing to say each item had an alpha of .7, since alpha pertains to the collection of items.)

Best wishes,
Michael T. Warren

References

1. Bardi A, Schwartz SH: Values and behavior: strength and structure of relations.*Pers Soc Psychol Bull*. 2003; **29** (10): 1207-20 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Developmental Psychology, Moral Development, Adolescent Development, Mindfulness

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 16 Jul 2020

Hyemin Han, University of Alabama, Tuscaloosa, USA

Dear Dr. Warren,

We sincerely appreciate your invaluable comments once again. We were able to improve the quality of our manuscript thanks to your constructive comments and suggestions. Here are our responses to your new comments:

1. I do remain skeptical about the word choice "improve your morals" found in the scale's items. My concern is that the phrase introduces unnecessary ambiguity. Are participants meant to interpret it as "improve yourself on specific moral traits" (based on your response I believe this is the intention), or are they meant to interpret it as "improve your moral values." To me, the former is directly relevant to becoming a better person, whereas the latter has very little to do with it-- we know, for example, that moral values and behaviour often show only small-to-moderate associations (e.g., Bardi & Schwartz, 20031). Becoming a better person entails not only holding moral values but embodying the cognitive, affective, motivational, and behavioral expressions of one's moral values. That said, it's entirely possible that I'm overthinking this and that typical survey respondents will intuitively grasp the meaning you intend. I'll leave it to you to decide whether this issue warrants further discussion.

Response: Thank you very much for your concern regarding the use of the terms once again. We

agree with you that the issue that you mentioned would be problematic in the long term, so we hope the issue is examined in our follow-up studies focusing on how different people differently perceive the concepts implied in the terms. In the revised manuscript, we added the point that you mentioned with additional citations:

There might be another concern regarding use of the terms, "morals and characters," in our measure related to whether a simple belief about the possibility to improve ones' morals (or moral values) is directly relevant to their moral growth. For instance, one previous study about the relationship between one's endorsed personal values and behavioral outcomes reported that the correlation between values and behaviors, particularly other-reported behaviors, was moderate at the greatest 40. Despite of this potential issue, however, we decided to use the terms because we tried to design the measure so that participants intuitively perceive and interpret items about "morals and characters" while considering their own beliefs about moral growth. In fact, researchers in implicit theories proposed that both incremental and entity theories are intuitive and "implicit" to people 41, so we intended to develop our items based on this point.

2. Beyond that, I would ask you to consider the following additional minor suggestions:Consider providing the MGM scale anchors either in the "Translation of the MGM measure to English" section (p. 3) or in a note beneath Table 2. I think requiring people to find the scale anchors in the "Supplementary Materials.docx" file may prove inconvenient; some future researchers may take the items from Table 2 and make up their own anchors.

Response: We appreciate your suggestion. In the revised manuscript, we presented the Likert scale that we used in the measure:

As a result, the tested measure included six items as well (e.g., "No matter who you are, you can significantly improve your morals and character") and answers were anchored to a six-point Likert scale (i.e., strongly disagree (1), disagree (2), mostly disagree (3), mostly agree (4), agree (5), strongly agree (6)) (see Extended data for the full measure 10).

3. In discussing Item 4's smaller factor loading in Study 2, I don't think it's fair to say it had a "slightly lower factor loading in Study 2 compared with Study 1." A difference of -.73 vs. -.39 in standardized loadings is quite substantial. Additionally, I wonder if readers will know which item is #4, since you're referring to the second item in Table 2. This issue also pertains to other items mentioned by number throughout the document.

Response: Thank you very much for your comment on the factor loading issue. In the revised manuscript, we added the point that you mentioned:

It would be an issue since researchers have regarded .40 as the threshold for a good factor loading 42.

In addition, although it is somehow preliminary at this point, we did a quick analysis of the data collected from a part of our follow-up studies ($N = 701$). The quick analysis showed that all factor loadings exceeded .75, so the aforementioned issue might be able to be addressed with further analysis with a larger dataset. Related to the policy of *F1000Research*, one good point to publish a paper in this platform is that it allows authors to post updated versions even after their paper has been approved by reviewers. So, to get the benefit from the platform policy that allows

post-publication updates, we decided to briefly report the preliminary result in the revised manuscript:

Although the issue could not be resolved completely with the current dataset, one larger dataset (N = 701 as of May 2020) that is currently being collected for the next research project was analyzed as a possible way to address the issue. When we conducted preliminary CFA with the new dataset, all four factor loadings were greater than .75 while the CFA model showed good model fit, RMSEA = .08, SRMR = .01, TLI = .98, CFI = .99. Given this, the small factor loading of Item 4 reported in Study 2, -.39, could be addressed in the long term with additional data collection and analysis.

4. Question: Is the Implicit Theory Measure in fact domain-general (as stated on p. 5) or does it measure intelligence growth mindset?

Response: We appreciate your comment that requests the clarification of the term. In the revised manuscript, we addressed the point:

We employed the Implicit Theory Measure 1 , which measures growth mindset in general, particularly intelligence growth mindset, and constitutes the basis of the MGM measure, to test convergent and discriminant validity.

5. I'm not sure what the following sentence means in the context of the model fit paragraph on p. 5: "As shown in Table 1, when we recalculated indices after exclusion of the items, they all remained greater than .7." I assume you're referring to alpha, but the context suggests you're saying the model fit indices were greater than .7, which is confusing. (It would also be confusing to say each item had an alpha of .7, since alpha pertains to the collection of items.)

Response: Thanks a lot for your request for the clarification of the concept. In the revised manuscript, we addressed the concern:

In addition, as shown in Table 1, when we recalculated reliability indices, Cronbach α and test-retest r, after exclusion of the items, they all remained greater than .7.

Again, we sincerely appreciate your comments and suggestions. They have significantly contributed to the improvement of the paper that we have been working on.

Best,
Hyemin

Competing Interests: N/A

Version 1

Reviewer Report 29 April 2020

?

Susan Mangan

[1] Department of Psychology, Claremont Graduate University, Claremont, USA

[2] Thrive Center for Human Development, Pasadena, California, USA

Han *et al.* suggest the creation of a moral growth mindset scale, which could make a significant contribution to the growing literature on growth mindset. The initial validity and reliability of this scale appears sound, and the inclusion of moral growth mindset (MGM) to the literature base seems evident. Despite the promise of this scale and the concept it captures, the article, as written, has some room for improvement. Most notably, the items themselves may benefit from further revision, more information is needed, especially on the various types of validity discussed in the article (e.g. convergent, divergent), and the writing could be clearer and more concise throughout.

Detailed responses:
- Lit review
  - One piece missing here is an argument for the significance of this scale. Why is moral growth mindset an important concept to measure? How does this measure add to existing literature (and existing measures on growth mindset + morality?) Were items for this scale taken from related scales (growth mindset?)

- Study 1
  - Elaborate on the "Implicit theory measure." Be sure to say 1-2 sentences about this to help us understand its significance to the current study.
  - *Participants*
    - How were participants compensated? Class credit? Gift card? No compensation?
    - (Ok, later in the next paragraph you mention class credit – I would mention in the first paragraph that participants were offered class credit to participate).
  - To help clarify each portion of the study and make it easy for readers to find the information that wish, I would add a heading of "Procedure" that begins with the "Participants received a link to the qulatircs survey…"
  - *Analysis*
    - What is "underlying data" and where is it located? From the final sentence of your analysis section.

- Study 2
  - "In study 2, we tested THE correlation between…" (Add "the" to sentence).
  - You mention the SONA system and how participants selected studies here – make sure to also include this recruitment information in study 1. Additionally, here you include information about the order of survey scales and demographics, which is also not in study 1. In general, there is different information here than in study 1 – I would align these participant sections to include the same relevant information. I would also, again, separate this into a clear "participants" section and a clear "procedure" section.

- Additionally: How long did it take participants to complete surveys? Were any attention check items included?
- *Measures*
  - While these measures look great – we need to understand their inclusion. In the literature review portion of this article there should be discussion of each construct and why/how you expect it to relate to moral growth mindset. Why are these good choices for convergent validity?
  - You mention that further details are available in "extended data" but I don't see a way to access this. Is all the relevant information explaining the connection of these scales to moral growth mindset provided there?
  - Additionally, did you test for divergent validity? Content validity? Construct validity? If not, why?
  - Ok, some indication of a test for discriminant validity is presented in analysis, but this should be listed with the other measures above. Similarly, more information is needed here about why you expect this to be divergent from moral growth mindset.

- Discussion
  - 1st sentence: "We developed and tested….from youth participants." These participants were emerging adults, correct? Not youth?
  - I would elaborate heavily on this first paragraph. What does it mean that moral disengagement was negatively correlated with MGM? What does it indicate if MGM and growth mindset were discriminant? As this is a measure development paper, I would include a richer discussion of what you found and what it indicates for each type of validity.
  - I'm not sure what you mean here in sentence two "In fact, the previous studies that developed and tested measurements for the mindset with diverse domains…" What is "the mindset"? Do you mean, that tested other types of domain-specific growth mindset?
  - There is a lot of information about previous studies coming up here that should also be in the introduction/literature review section. This would be great to include so we know what you are expecting before the study is run. Then, here in the discussion, you can tell us if your predictions came out as expected or not.
  - Paragraph 1 and 2 here are a little confusing. Be sure to start each paragraph with a general sentence that indicates what you found. Then, discuss each of your results that provides evidence for your statement, and conclude with a sentence that indicates to us the importance of these results. It feels like there are too many concepts included in each paragraph, and the writing is a bit confusing throughout.
  - Last paragraph "However, there are limitationS" (missing s).

- Items
  - These items all feel a bit too similar – Each asks if morality can or cannot be "improved." Did you consider other items with different word choices? For example, even "You can always become a more moral person with better character." Or "It is possible to grow in your character and morality." These items are so similar that it feels hard to argue any differences between them.
  - Additionally, if possible it's always best to include simpler words over more complicated ones. For instance "substantially" could easily be "a lot" and "considerably" could also be "a lot." More complex words tax the participants a bit more, and may make the sentences more difficult to digest, especially for those with lower readings levels or from less educated backgrounds.

- General notes
  - There are a few places where the writing could be condensed or the use of alternate terms would improve ease of reading. For instance, an example of condensing would be: in participants, "Participants were recruited from an undergraduate subject pool. The pool consisted of students who were enrolled in educational psychology classes" could be condensed to "Participants were recruited from students enrolled in educational psychology classes." A for instance of somewhere where alternate terms would be benefitial would be in: Results, "First, all consistency indicators indicated…" Instead of using "indicator/indicated" here I would recommend "First, all consistency indicators revealed…" Or, in the spirit of condensing/being more specific, I might suggest: First, the measure demonstrated at least acceptable reliability according to both cronbach's alpha values and test-retest reliability."
    - Look for these types of instances throughout the paper with an eye towards condensing repetitive language and becoming more specific.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Partly

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Developmental psychology, positive psychology, emerging adulthood, positive pyschology interventions

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 06 May 2020

Hyemin Han, University of Alabama, Tuscaloosa, USA

Dear Dr. Mangan,

We sincerely appreciate your comments and suggestions on our manuscript. We found that they

are very constructive and information. While revising our manuscript, we have done our best to address the concerns that you mentioned in your review report. Please find our responses to your comments below. Thank you very much for your time and consideration.

Best,
Hyemin Han (Corresponding author)

Responses

1. One piece missing here is an argument for the significance of this scale. Why is moral growth mindset an important concept to measure? How does this measure add to existing literature (and existing measures on growth mindset + morality?) Were items for this scale taken from related scales (growth mindset?)

Response: We appreciate your comment regarding the explanation of why the measure is important. Also, we described how the items were developed. In the revised manuscript, we elaborated further details.

"The results suggested that among younger populations, MGM might increase participants' prosocial behavior due to the belief that it will make them morally better. Given this, MGM would be considered as a factor that contributes to moral development. In order to adequately examine how MGM contributes to moral development, however, it is necessary to have an appropriate measure. Additionally, if moral growth mindset motivates people to learn how to become more moral, as previous research suggests, then it is important for moral educators to have a tool to assess the malleability beliefs students have related to their morals. For example, if moral educators are able to identify that some students have a fixed mindset related to their morals, then an appropriate starting point may be to provide them with evidence that it is possible to improve moral character throughout one's life."

"Instead, the inventors (HH, KJD, and YJC) of the Korean MGM measure created its English version based on the structure of the Korean version and the wording in the Implicit Theory measure. In addition, the Implicit Theory measure was used due to the fact that it had six items and was based on Dweck's original measure of growth mindset for intelligence. As a result, the tested measure included six items as well (e.g., "No matter who you are, you can significantly improve your morals and character") and answers were anchored to a six-point Likert scale (see Extended data for the full measure 10)."

2. Elaborate on the "Implicit theory measure." Be sure to say 1-2 sentences about this to help us understand its significance to the current study.

Response: Thank you very much for your suggestion regarding the elaboration of the construct. We elaborated such a point in the revised manuscript:

"Growth mindset refers to the belief that it is possible to improve one's abilities and qualities, such as intelligence or personality 1 . These individuals believe that this can be done through effort and learning, which helps fosters motivation. Higher motivation for those with a growth mindset is encouraged through having attitudes such as viewing hardships as a chance to work harder rather

than an indication of failure, and striving for success due to genuinely wanting to learn instead of being concerned with how others view them 2"

3. How were participants compensated? Class credit? Gift card? No compensation?(Ok, later in the next paragraph you mention class credit – I would mention in the first paragraph that participants were offered class credit to participate).

Response: Thanks a lot for your request for the clarification of the compensation. In the revised manuscript, such a point is more clearly stated:

"Participants were recruited from students enrolled in undergraduate educational psychology classes and they were provided with a course credit."

4. To help clarify each portion of the study and make it easy for readers to find the information that wish, I would add a heading of "Procedure" that begins with the "Participants received a link to the qulatircs survey…"

Response: We sincerely appreciate your suggestion regarding the use of the "procedure" subsection. In the revised manuscript, we created the new subsection for a better structure.

5. What is "underlying data" and where is it located? From the final sentence of your analysis section.

Response: Thank you for your comment regarding "underlying data." "Underlying data" is a way to include supplementary materials in F1000Research. Readers can download the supplementary materials with the URL provided at the end of the main text. More specifically, a link to an open science repository is provided in the "data statement" section as per the journal guidelines.

6. "In study 2, we tested THE correlation between…" (Add "the" to sentence).You mention the SONA system and how participants selected studies here – make sure to also include this recruitment information in study 1. Additionally, here you include information about the order of survey scales and demographics, which is also not in study 1. In general, there is different information here than in study 1 – I would align these participant sections to include the same relevant information. I would also, again, separate this into a clear "participants" section and a clear "procedure" section.

Response: We appreciate your comments regarding the typo and the use of the independent subsection, "procedure." We addressed these issues in the revised manuscript.

7. Additionally: How long did it take participants to complete surveys? Were any attention check items included?

Response: Thanks a lot for your kind comment regarding the survey duration. Unfortunately, we could not include any attention check items in the survey form. We explained further details in the

limitation section:

"Third, although participants spent about 33.98 minutes (median) to complete Study 2 we did not include any attention check items."

8. While these measures look great – we need to understand their inclusion. In the literature review portion of this article there should be discussion of each construct and why/how you expect it to relate to moral growth mindset. Why are these good choices for convergent validity?

Response: We sincerely appreciate your comment regarding the rationale for the inclusion of the additional measures. In the introduction section in Study 2, we added explanations regarding the point. In addition, while describing each additional measure, we explained the rational for the inclusion as well as the hypothesized correlation with MGM.

"We selected several moral and positive psychological measures to test the convergent and divergent validity of the MGM measure. We employed the Implicit Theory Measure 1, which measures domain-general growth mindset and constitutes the basis of the MGM measure, to test convergent and discriminant validity. For the selection of moral psychological measures, we referred to recent articles about psychological constructs that significantly predict prosocial and civic behavior 31. They proposed moral judgment 18, 19, moral emotion (empathy) 20, and moral identity 21 as fundamental constructs in moral functioning. We also employed the Propensity to Morally Disengage Scale to examine whether the MGM showed negative correlation with moral disengagement 22 since Han et al. (2018) 4 reported that MGM promotes moral engagement. In addition to the aforementioned moral psychological measures, we used the Claremont Purpose Scale as a way to examine one's positive development in terms of flourishing 23, given that purpose has been regarded as a possible moral virtue for eudemonic wellbeing 32.
In general, according to the previous studies that examined the relationship between growth mindset, positive psychological indicators, and antisocial tendency (e.g., 24– 26), we hypothesized that the sizes of correlation coefficients between MGM and other indicators, except the general growth mindset, would be between .10 (small) and .30 (medium). We discussed further details regarding the hypothesized effect size of each measure in the following sections."

9. You mention that further details are available in "extended data" but I don't see a way to access this. Is all the relevant information explaining the connection of these scales to moral growth mindset provided there?

Response: Thank you for your comment regarding the "extended data." Same to the "underlying data," "extended data" can also be downloaded with the link provided at the end of the main text. In the revised manuscript, as we mentioned in our response to your comment above, we explained further details regarding each additional measure in the methods section in Study 2.

10. Additionally, did you test for divergent validity? Content validity? Construct validity? If not, why? Ok, some indication of a test for discriminant validity is presented in analysis, but this should be listed with the other measures above. Similarly, more information is needed here about why you expect this to be divergent from moral growth mindset.

Response: We appreciate your comment regarding the clarification of the tests that we conducted. In the revised manuscript, we explained further details.

"We aimed at testing the validity of the measure, construct, convergent, and divergent validity. We selected several moral and positive psychological measures to test the convergent and divergent validity of the MGM measure. We employed the Implicit Theory Measure 1, which measures domain-general growth mindset and constitutes the basis of the MGM measure, to test convergent and discriminant validity. For the selection of moral psychological measures, we referred to recent articles about psychological constructs that significantly predict prosocial and civic behavior 31. They proposed moral judgment 18, 19, moral emotion (empathy) 20, and moral identity 21 as fundamental constructs in moral functioning. We also employed the Propensity to Morally Disengage Scale to examine whether the MGM showed negative correlation with moral disengagement 22 since Han et al. (2018) 4 reported that MGM promotes moral engagement. In addition to the aforementioned moral psychological measures, we used the Claremont Purpose Scale as a way to examine one's positive development in terms of flourishing 23, given that purpose has been regarded as a possible moral virtue for eudemonic wellbeing 32."

"Given that the Implicit Theory Measure measures one's general growth mindset, we expected that it would be positively correlated with MGM. However, because the construct measured by the Implicit Theory Measure is not domain specific, we also expected that the MGM would not completely overlap with this construct (discriminant validity). Given these, the effect size of the correlation coefficient would be medium to large (r = +.3 - +.5)."

11. 1st sentence: "We developed and tested….from youth participants." These participants were emerging adults, correct? Not youth?

Response: Thanks for your point regarding the correct use of the term. Yes, that is correct. So, we used "emerging adults" instead of "youth…" in the revised manuscript:

"We developed and tested the English version of the MGM measure in this study with data collected from emerging adult participants."

12. I would elaborate heavily on this first paragraph. What does it mean that moral disengagement was negatively correlated with MGM? What does it indicate if MGM and growth mindset were discriminant? As this is a measure development paper, I would include a richer discussion of what you found and what it indicates for each type of validity.

Response: We appreciate your comment regarding moral disengagement. In the section that describes moral disengagement and its measure, we address the points that you mentioned:

"Propensity to Morally Disengage Scale. The moral disengagement scale measures one's propensity to disengage from moral behavior within morally problematic situations 22. It measures moral disengagement propensities for eight mechanisms (i.e., moral justification, euphemistic labeling, advantageous comparison, displacement of responsibility, diffusion of responsibility, distortion of consequences, dehumanization, attribution of blame) with eight items (one item per mechanism). We used a composite score of the eight items. The internal structure of the scale was tested with CFA by Moore et al. (2012) 22. As Bandura (2002) proposed 35, moral disengagement

is negatively associated with motivation for moral engagement. Thus, we expected moral disengagement would be negatively associated with MGM while the effect size of the correlation would be similar to the cases of the IRI and MIS (small to medium; r = -.1 - .3)."

In addition, in the discussion section, we elaborated the meaning of the negative correlation found from our analysis:

"In addition, moral disengagement was negatively correlated with MGM. Since moral disengagement allows people to dismiss negative feelings, they may have about behaving immorally using the eight mechanisms previously mentioned, this increases the likelihood of continuing to behave immorally. In this way, moral disengagement and MGM have somewhat reverse trajectories. As hypothesized, this suggests that MGM may promote engaging in moral behavior. In addition, since moral internalization, which has been shown to inhibit moral disengagement 39, was also positively correlated with MGM, it makes sense that our measure was negatively correlated with moral disengagement. If somebody has a strong sense of their morals and these values are internalized, this may help them to stay engaged with their standards and furthermore, be motivated to continue to be morally better."

13. I'm not sure what you mean here in sentence two "In fact, the previous studies that developed and tested measurements for the mindset with diverse domains…" What is "the mindset"? Do you mean, that tested other types of domain-specific growth mindset?

Response: Thank you for your request for the clarification. Yes, that is correct. In the revised manuscript, we specified the nature of the mindset:

"In fact, the previous studies that developed and tested measurements for diverse types of domain-specific growth mindset have shown that the measurements possessed good reliability and validity as well (e.g., 29, 30)."

14. There is a lot of information about previous studies coming up here that should also be in the introduction/literature review section. This would be great to include so we know what you are expecting before the study is run. Then, here in the discussion, you can tell us if your predictions came out as expected or not.

Response: We sincerely appreciate your suggestion regarding rewriting the introduction. We agree with you that some theoretical contents that were presented in the discussion section in the original manuscript could be moved on to the introduction section for a better structure. Following your suggestion, in the revised manuscript, we presented such contents in the general introduction or introduction of each study.

15. Paragraph 1 and 2 here are a little confusing. Be sure to start each paragraph with a general sentence that indicates what you found. Then, discuss each of your results that provides evidence for your statement, and conclude with a sentence that indicates to us the importance of these results. It feels like there are too many concepts included in each paragraph, and the writing is a bit confusing throughout.

Response: We appreciate your comment regarding the discussion section. As you suggested, in the revised manuscript, we slightly restructured the paragraphs in the discussion section. We revised each of following paragraphs so that it discusses one specific point each time.

16. Last paragraph "However, there are limitationS" (missing s).

Response: Thank you very much for your comment on the typo. We corrected the point in the revised manuscript.

17. These items all feel a bit too similar – Each asks if morality can or cannot be "improved." Did you consider other items with different word choices? For example, even "You can always become a more moral person with better character." Or "It is possible to grow in your character and morality." These items are so similar that it feels hard to argue any differences between them. Additionally, if possible it's always best to include simpler words over more complicated ones. For instance "substantially" could easily be "a lot" and "considerably" could also be "a lot." More complex words tax the participants a bit more, and may make the sentences more difficult to digest, especially for those with lower readings levels or from less educated backgrounds.

Response: We appreciate your comments regarding the items used in our measure. Yes, we agree with you that some words used in the items are somehow complex to be easily understood by younger participants. So, we also think that such items may need to be modified if the measure is to be administrated among younger populations. Also, we also acknowledged that some words (e.g., "improve") were repeatedly used in multiple items. Since we intended to keep the consistency with the original measures that we referred to (e.g., Dweck's general growth mindset measure, the Korean version of the MGM measure), we ended up with using such terms in our measure. We explained these points in the limitation section:

"Fifth, the items used in the MGM measure could be revised particularly when being administered among younger populations. We decided to use the current wordings to maintain consistency with the Korean version of the MGM measure and the Implicit Theory Measure, which constituted the basis of our measure. However, to make the measure more applicable to younger populations, some complex words (e.g., "substantially," "considerably") could be replaced with simpler words (e.g., "a lot"). Finally, since several items in the measure might seem to be similar, the words could be revised in future studies, particularly those focusing on children or young adolescents."

18. There are a few places where the writing could be condensed or the use of alternate terms would improve ease of reading. For instance, an example of condensing would be: in participants, "Participants were recruited from an undergraduate subject pool. The pool consisted of students who were enrolled in educational psychology classes" could be condensed to "Participants were recruited from students enrolled in educational psychology classes." A for instance of somewhere where alternate terms would be benefitial would be in: Results, "First, all consistency indicators indicated…" Instead of using "indicator/indicated" here I would recommend "First, all consistency indicators revealed…" Or, in the spirit of condensing/being more specific, I might suggest: First, the measure demonstrated at least acceptable reliability according to both cronbach's alpha values and test-retest reliability."
Look for these types of instances throughout the paper with an eye towards condensing repetitive

language and becoming more specific.

Response: We sincerely appreciate your suggestions regarding the brevity of the manuscript. We agree with you that increasing the brevity is essential to enable potential readers to better understand the overall theme of our manuscript while saving their time. Thus, we edited the whole manuscript during the current revision process. In addition, we revised the manuscript as an effort to minimize repeating to use the same words in the multiple places.

Competing Interests: Not available.

Reviewer Report 27 April 2020

https://doi.org/10.5256/f1000research.25564.r62240

Michael T. Warren [iD]

[1] Human Early Learning Partnership, University of British Columbia, British Columbia, Canada
[2] Psychology Department, Western Washington University, Bellingham, WA, USA

Han and colleagues examined the psychometric properties of the English version of a growth mindset measure in the moral domain. Moral growth mindset (MGM) may prove to be a useful motivational construct in the scientific study of morality in general, and moral development in particular. I believe the authors present sufficient initial evidence of the validity and reliability of their measure, marking an important step in the scientific study of MGM in English-speaking populations.

I see three major issues with the paper in its current form, and I would encourage the authors to revise their manuscript in light of my suggestions.

First, my biggest concern is the use of phrases such as "improve one's morals" and "improve your morals." The former occurs in the paper's conceptual framing (p. 3) and the latter appears in each of the MGM scale's four items (p. 4). The issue with these phrases is that improving one's morals seems to deviate conceptually from improving one's morality. One might improve their morals by setting new (or higher) moral standards for themselves, yet they may fail miserably in living up to their moral values. By contrast, improving one's morality involves actually becoming a better person, and this, I believe, is the construct the authors intended to measure. Since in my view the items miss the target to some extent, I have indicated that the work is only "partly" technically sound. Unfortunately, I don't think much can be done about this issue at this point, but at a minimum I would recommend that the authors either provide an argument for the use of "moral" rather than "morality" in their scale, or identify this as a limitation of their scale. In addition, they might choose to argue that this concern is assuaged by the scale's strong evidence of convergent validity with other measures of morality.

Second, I think the CFA results need to be communicated more fully, and that is why I have indicated that the statistical analyses and their interpretations are only "partly" appropriate. On p. 4, I certainly understand consulting previous studies (e.g., Han *et al.*, 2018), but data from the current (i.e., English)

study should be given primary importance in refining the English scale. I recommend reporting the factor loadings from the original CFA (i.e., before Items 1 and 2 were removed), so readers can evaluate whether removing these items was justified on empirical grounds. On a related note, I think it would be appropriate to acknowledge the small factor loading (-.39) for the reverse-scored item in Study 2.

Third, I think the introduction to Study 2 (p. 5) should be expanded considerably. It would be very helpful to provide a brief rationale for why the selected constructs were chosen for convergent and discriminant validity testing. In addition, it would be helpful to specify hypotheses concerning the strength and direction of the associations between MGM and other constructs (i.e., with which constructs does MGM have strongest and weakest theoretical ties?), and why. The discussion currently states that the observed associations were "as hypothesized," but no hypotheses were specified in the lead-up to Study 2. I also found myself wondering why Chiu *et al.*[1] (1997)'s original 3-item English MGM measure was not included for convergent and incremental validity testing.

Minor comments:
- Page 3: My understanding is that growth mindset generally concerns one's beliefs about the malleability of one's own (and others') qualities. Thus, it seems a little bit too generic to define growth mindset as believing "it is possible to improve aspects of one's life." There are aspects of a person's life (e.g., what kind of work they do; where they live, etc.) that are not qualities of their personhood. I suggest the authors consider revising their opening definition of growth mindset.

- Page 3: It's not clear to me how allowing participants to define "moral" and "character" necessarily allows them to do so "without bias." Instead, I think it would be more accurate to say that the approach taken leaves it up to participants to interpret "moral" and "character" according to their own subjective understandings of those terms. (Note that this approach makes no claim that participants' understandings are "without bias.")

- Pages 3 and 5: I suggest the authors change the "Participants" heading to "Participants and procedures."

- Page 4: I suggest confirming that three IRB approvals were needed for just two studies.

- Pages 4-5: I suggest referring to model fit and reliability indices as either "indices" or "indexes," rather than "indicators." Given that CFA was involved, readers may assume "indicators" refers to measured variables loading onto latent factors.

- Page 5 (last paragraph of Study 1): I would like to suggest an alternative explanation as to why Items 1 and 2 (presumably) had lower factor loadings. These two were the only items to convey morality/character as dispositional (e.g., "You have a certain morality and character..."; "Your morality and character are something about you..."). By contrast, all items measured malleability beliefs, including the retained reverse-scored item ("To be honest, you can't really improve your morals and character."). My understanding is that a growth mindset is anchored in malleability beliefs, and having a growth mindset does not preclude the belief in moral dispositions (e.g., with effort I can become a more consistently/dispositionally honest person). In other words, perhaps the reason why Items 1 and 2 presumably had lower factor loadings was because they strayed somewhat from the core of the growth mindset construct (i.e., malleability beliefs), rather than because they used the vague qualifier, "much." Just some food for thought.

- Page 5 (Participants section): Much of the first two paragraphs in this section is redundant with the procedures described in Study 1. The authors may wish to simply state that the same recruitment procedures were used as described in Study 1.

- Page 5: I would strongly urge the authors to omit the term, "marginally correlated" in relation to MGM's association with the bDIT. Once a threshold for statistical significance has been set (e.g., .05), a finding is either statistically significant or non-significant. Correlations with p-values between .05 and .10 are non-significant.

- Page 6 (Table 3): I suggest indicating where Cronbach alphas are reported (i.e., on the diagonal).

- Page 6: More information on the potential utility of the MGM measure for understanding moral development would be a nice selling point for the scale. For example, this scale makes it possible to test whether MGM moderates the efficacy of moral education and social emotional learning interventions. The scale would also be an important outcome measure in examining how to nurture MGM (e.g., through process praise, teaching about neuroplasticity, etc.).

- Page 6: It is not yet clear why the authors would like to have conducted CFAs for the other measures. I would suggest they either drop this piece or further explain why additional CFAs would be desirable if the sample were large enough.

- Page 6: I think more explanation is needed as to why testing measurement invariance would be helpful. For example, the authors might say that examining measurement invariance across diverse groups of people (e.g., political conservatives vs. liberals; young adults vs. older adults) would help evaluate whether the scale—which leaves it up to participants to interpret morality and character in the item stems—in fact measures the same thing for groups who may use different underlying folk conceptions of morality.

References
1. Chiu C, Dweck C, Tong J, Fu J: Implicit theories and conceptions of morality.*Journal of Personality and Social Psychology*. 1997; **73** (5): 923-940 Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Partly

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Partly

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Developmental Psychology, Moral Development

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 06 May 2020

Hyemin Han, University of Alabama, Tuscaloosa, USA

Dear Dr. Warren,

We sincerely appreciate your comments and suggestions on our manuscript. We found that they are very constructive and information. While revising our manuscript, we have done our best to address the concerns that you mentioned in your review report. Please find our responses to your comments below. Thank you very much for your time and consideration.

Best,
Hyemin Han (Corresponding author)

Responses

1. First, my biggest concern is the use of phrases such as "improve one's morals" and "improve your morals." The former occurs in the paper's conceptual framing (p. 3) and the latter appears in each of the MGM scale's four items (p. 4). The issue with these phrases is that improving one's morals seems to deviate conceptually from improving one's morality. One might improve their morals by setting new (or higher) moral standards for themselves, yet they may fail miserably in living up to their moral values. By contrast, improving one's morality involves actually becoming a better person, and this, I believe, is the construct the authors intended to measure. Since in my view the items miss the target to some extent, I have indicated that the work is only "partly" technically sound. Unfortunately, I don't think much can be done about this issue at this point, but at a minimum I would recommend that the authors either provide an argument for the use of "moral" rather than "morality" in their scale, or identify this as a limitation of their scale. In addition, they might choose to argue that this concern is assuaged by the scale's strong evidence of convergent validity with other measures of morality.

Response: Thank you very much for your comment regarding the use of the terms in our study. In the revised manuscript, we added a paragraph describing why we decided to use "morals" instead of "morality" in our measure. The point is that we intended use the term to maintain the consistency with the prior study.

"Finally, since Chiu et al. used terms related to specific morals and characteristics in their original three-item subscale (e.g., "A person's moral character," "whether a person is responsible and sincere," "a person's moral traits"), we decided to use "morals and character" in order to stay consistent with the construct they were measuring. That is, rather than measuring participants'

malleability beliefs about the overarching system of values they have, we wanted to measure malleability beliefs regarding individual morals, as did the original measure. Doing so may increase the chance for interventions since if people want to become a better person (improve their morality) they may need to believe that their values (morals) can be improved."

2. Second, I think the CFA results need to be communicated more fully, and that is why I have indicated that the statistical analyses and their interpretations are only "partly" appropriate. On p. 4, I certainly understand consulting previous studies (e.g., Han *et al.*, 2018), but data from the current (i.e., English) study should be given primary importance in refining the English scale. I recommend reporting the factor loadings from the original CFA (i.e., before Items 1 and 2 were removed), so readers can evaluate whether removing these items was justified on empirical grounds. On a related note, I think it would be appropriate to acknowledge the small factor loading (-.39) for the reverse-scored item in Study 2.

Response: We appreciate your comment regarding how to report results from CFA. Following your suggestion, we added a supplementary table that demonstrates the factor loadings in 6-item and 5-item models. As you can see, Item 1 and Item 2 showed the lowest standardized factor loadings in the 6-item and 5-item models, respectively. In the revised manuscript, we mentioned the point that they were excluded from the measure due to their lowest standardized factor loadings.

Table S1
Factor loadings from the CFA of the six- and five-item models in Study 1

Six-item model
Five-item model

Item
Unstandardized
Standardized
Unstandardized
Standardized

You have certain morals and character, and you can't really do much to improve it.
-.71
-.58
-
-

Your morals and character are something about you that you can't improve very much.
.70
-.67
-.61
-.58

No matter who you are, you can significantly improve your morals and character.
.64
.62

.68
.65

To be honest, you can't really improve your morals and character.
-.86
-.86
-.81
-.81

You can always substantially improve your morals and character.
.66
.69
.72
.75

You can improve your basic morals and character considerably.
.80
.'.83
.85
.88

Note: Bolded items (Items 1 and 2) were excluded from the finalized MGM measure.

"In the supplementary table in Underlying data, we presented factor loadings for the six-item and five-item models. In the six-item model, Item 1 showed the lowest standardized factor loading, identical to what was reported in Han et al. (2018) 4. After excluding Item 1, Item 2 showed the lowest standardized loading in the five-item model, so we removed this item accordingly."

Moreover, we acknowledge the slightly low factor loading in Study 2:

"However, it should be acknowledged that Item 4 showed a slightly lower factor loading in Study 2 compared with Study 1, although the overall model fit indices were excellent. This point might need to be tested in future studies with more samples."

3. Third, I think the introduction to Study 2 (p. 5) should be expanded considerably. It would be very helpful to provide a brief rationale for why the selected constructs were chosen for convergent and discriminant validity testing. In addition, it would be helpful to specify hypotheses concerning the strength and direction of the associations between MGM and other constructs (i.e., with which constructs does MGM have strongest and weakest theoretical ties?), and why. The discussion currently states that the observed associations were "as hypothesized," but no hypotheses were specified in the lead-up to Study 2. I also found myself wondering why Chiu *et al.* (1997)'s original 3-item English MGM measure was not included for convergent and incremental validity testing.

Response: Thanks a lot for your suggestion regarding the expansion of the Study 2 introduction. Following your suggestion, in the revised manuscript, we elaborated the rationale regarding how the additional construct used in our study were selected with citations. Furthermore, in the methods section, per additional measurement, we explained the direction and effect size of the hypothesized correlation.

"We selected several moral and positive psychological measures to test the convergent and divergent validity of the MGM measure. We employed the Implicit Theory Measure 1, which measures domain-general growth mindset and constitutes the basis of the MGM measure, to test convergent and discriminant validity. For the selection of moral psychological measures, we referred to recent articles about psychological constructs that significantly predict prosocial and civic behavior 31. They proposed moral judgment 18, 19, moral emotion (empathy) 20, and moral identity 21 as fundamental constructs in moral functioning. We also employed the Propensity to Morally Disengage Scale to examine whether the MGM showed negative correlation with moral disengagement 22 since Han et al. (2018) 4 reported that MGM promotes moral engagement. In addition to the aforementioned moral psychological measures, we used the Claremont Purpose Scale as a way to examine one's positive development in terms of flourishing 23, given that purpose has been regarded as a possible moral virtue for eudemonic wellbeing 32.
In general, according to the previous studies that examined the relationship between growth mindset, positive psychological indicators, and antisocial tendency (e.g., 24– 26), we hypothesized that the sizes of correlation coefficients between MGM and other indicators, except the general growth mindset, would be between .10 (small) and .30 (medium). We discussed further details regarding the hypothesized effect size of each measure in the following sections."

We agree with you that employing Chiu et al.'s original measure in the present study would be beneficial. However, we did not consider doing so because our measure was originally Based on Dweck's updated six-item measure for general growth mindset. In the limitation section, we acknowledged your point for reader's information.

"Fourth, we did not employ Chiu et al.'s (1997) 5 original measure, which could be informative while conducting the convergent validity check, although our measure was based on Dweck's (2000) 1 updated six-item general growth mindset measure."

4. Page 3: My understanding is that growth mindset generally concerns one's beliefs about the malleability of one's own (and others') qualities. Thus, it seems a little bit too generic to define growth mindset as believing "it is possible to improve aspects of one's life." There are aspects of a person's life (e.g., what kind of work they do; where they live, etc.) that are not qualities of their personhood. I suggest the authors consider revising their opening definition of growth mindset.

Response: We appreciate your suggestion regarding the introduction. We revised the introduction for a better definition of the growth mindset:

"Growth mindset refers to the belief that it is possible to improve one's abilities and qualities, such as intelligence or personality 1 . These individuals believe that this can be done through effort and learning, which helps fosters motivation. Higher motivation for those with a growth mindset is encouraged through having attitudes such as viewing hardships as a chance to work harder rather than an indication of failure, and striving for success due to genuinely wanting to learn instead of being concerned with how others view them 2"

5. Page 3: It's not clear to me how allowing participants to define "moral" and "character" necessarily allows them to do so "without bias." Instead, I think it would be more accurate to say that the approach taken leaves it up to participants to interpret "moral" and "character" according to

their own subjective understandings of those terms. (Note that this approach makes no claim that participants' understandings are "without bias.")

Response: We appreciate your comment regarding the use of the terms in our study. In the revised manuscript, we updated our explanation regarding the terms as per your comment:

"Thus, we used "morals and characters" in order for participants to be able to define the terms based on their own experiences and understanding."

6. Pages 3 and 5: I suggest the authors change the "Participants" heading to "Participants and procedures."

Response: Thanks a lot for your suggestion regarding the subsection. In the revised manuscript, following your and Dr. Mangan's suggestions, we moved contents regarding the study procedures to a new subsection, "procedures."

7. Page 4: I suggest confirming that three IRB approvals were needed for just two studies.

Response: Thank you for your comment regarding the IRB numbers. In the revised manuscript, we clearly stated which IRB protocols are relevant to which specific study.

8. Pages 4-5: I suggest referring to model fit and reliability indices as either "indices" or "indexes," rather than "indicators." Given that CFA was involved, readers may assume "indicators" refers to measured variables loading onto latent factors.

Response: We appreciate your comment regarding the use of the term. In the revised manuscript, as per your comment, we used "indices" in lieu of "indicators" while addressing CFA.

9. Page 5 (last paragraph of Study 1): I would like to suggest an alternative explanation as to why Items 1 and 2 (presumably) had lower factor loadings. These two were the only items to convey morality/character as dispositional (e.g., "You have a certain morality and character..."; "Your morality and character are something about you..."). By contrast, all items measured malleability beliefs, including the retained reverse-scored item ("To be honest, you can't really improve your morals and character."). My understanding is that a growth mindset is anchored in malleability beliefs, and having a growth mindset does not preclude the belief in moral dispositions (e.g., with effort I can become a more consistently/dispositionally honest person). In other words, perhaps the reason why Items 1 and 2 presumably had lower factor loadings was because they strayed somewhat from the core of the growth mindset construct (i.e., malleability beliefs), rather than because they used the vague qualifier, "much." Just some food for thought.

Response: Thanks a lot for the alternative explanation of the lower factor loadings of items 1 and 2. We added such an alternative explanation in the revised manuscript for readers' information:

"In addition, as another possibility, items 1 and 2 are more likely about entity beliefs, not malleability beliefs that constitute the basis of growth mindset. These items contain some words

*perhaps related to entity beliefs (e.g., "certain morals and characters...," "something about you…"),
so they might not directly measure the core of the growth mindset construct and showed lower
factor loadings compared to the other items."*

10. Page 5 (Participants section): Much of the first two paragraphs in this section is redundant with
the procedures described in Study 1. The authors may wish to simply state that the same
recruitment procedures were used as described in Study 1.

Response: Thank you very much for your suggestion for the brevity of our manuscript. We
shortened the redundant part in Study 2 as per your suggestion.

11. Page 5: I would strongly urge the authors to omit the term, "marginally correlated" in relation to
MGM's association with the bDIT. Once a threshold for statistical significance has been set (e.g.,
.05), a finding is either statistically significant or non-significant. Correlations with p-values between
.05 and .10 are non-significant.

Response: Thanks a lot for your comment about the use of the term, "marginally correlated." We
agree with you that the use of the term is somehow inappropriate, so in the revised manuscript, we
changed the part about interpreting the finding from correlation analysis:

*"The effect size of the correlation coefficient between MGM and bDIT was small as predicted, but
the correlation was non-significant (p = .08)."*

12. Page 6 (Table 3): I suggest indicating where Cronbach alphas are reported (i.e., on the
diagonal).

Response: We appreciate your suggestion. We added a brief description about where alphas
were reported:

"Cronbach αs are also reported (on the diagonal)."

13. Page 6: More information on the potential utility of the MGM measure for understanding moral
development would be a nice selling point for the scale. For example, this scale makes it possible
to test whether MGM moderates the efficacy of moral education and social emotional learning
interventions. The scale would also be an important outcome measure in examining how to nurture
MGM (e.g., through process praise, teaching about neuroplasticity, etc.).

Response: Thanks a lot for your suggestion about the elaboration of the potential utility of the
measure. In the introduction, we briefly mentioned how the measure could be used in moral
education:

*"Additionally, if moral growth mindset motivates people to learn how to become more moral, as
previous research suggests, then it is important for moral educators to have a tool to assess the
malleability beliefs students have related to their morals. For example, if moral educators are able
to identify that some students have a fixed mindset related to their morals, then an appropriate*

starting point may be to provide them with evidence that it is possible to improve moral character throughout one's life."

14. Page 6: It is not yet clear why the authors would like to have conducted CFAs for the other measures. I would suggest they either drop this piece or further explain why additional CFAs would be desirable if the sample were large enough.

Response: We appreciate your comment regarding the CFA of additional measures. We dropped the part as per your comment since it was not essential in our study.

15. Page 6: I think more explanation is needed as to why testing measurement invariance would be helpful. For example, the authors might say that examining measurement invariance across diverse groups of people (e.g., political conservatives vs. liberals; young adults vs. older adults) would help evaluate whether the scale—which leaves it up to participants to interpret morality and character in the item stems—in fact measures the same thing for groups who may use different underlying folk conceptions of morality.

Response: Thank you very much for your suggestion. We clarified the point in the revised manuscript:

"To address this issue, measurement invariance test would be a way to examine whether the MGM measure, which allows participants to interpret "morals" and "character" by themselves, measures the same construct across different groups who may use different underlying folk conceptions of morals and character."

Competing Interests: Not available.

Comments on this article

Version 1

Author Response 17 Apr 2020

Hyemin Han, University of Alabama, Tuscaloosa, USA

I would like to add our responses to previous reviewers' comments for readers' information. This paper was submitted to another journal, but rejected after one round of major revision. Some reviewers evaluated our manuscript favorably, so we decided to revise the manuscript based on their comments to improve its quality. Here are our responses:

Responses to the reviewer's comments

1.   The points that I have raised were in great part addressed. Below I write a few minor points about how I think the revisions should be improved.

- Participation rate. I think that the paper should include the information that all the invited students participated, because they were required to do so to get a course credit.

We appreciate your comment regarding the participant rate. In the revised manuscript, we clearly described that all participations were done voluntarily and all of them who appropriately signed up for our study completed the survey.

Only the participants who voluntarily signed for Study 1 were provided with the link. We created our Qualtrics survey in a way so that only the participants who answered all survey questions were able to complete the survey and receive a credit. Thus, there was no missing data in the present study. Afterwards, we sent them invitations to participate in the survey again one week later.

2. Missing data. Even when missing data was not an issue, it would be a good idea to state the % of missing values in the text. I assume that the online survey did not force the participants to respond to each item. If yes, that if okay, but even in that case you could state the (0)% of missing values explicitly.

Thank you very much for your comment regarding the missing data. As we responded to your prior comment, in the revised manuscript, we explicitly mentioned that there was no missing data.

3. Limitations. Was the course/the pool related to moral development, social psychology or related issues? If yes, the comment about the limited generalizability should be elaborated a bit in this context as the sample might over-represent people interested in character development, human relations etc.

We appreciate your comment regarding the nature of the pools. All participants were taking general psychology and educational psychology classes. Although some class contents were related to human development in general, the classes did not focus on moral and social development. In the revised manuscript, we explained the nature of the pools briefly.

Participants were recruited from an undergraduate subject pool. The pool consisted of students who were enrolled in introductory psychology and educational psychology classes.

4. Discussion. I still find pieces of the discussion unelaborated. Specifically, as a reader I would like to see there the novel findings of the study interpreted in the context of the existing knowledge with a few citations of the existing literature.

Thanks a lot for your suggestion regarding the elaboration of the discussion section. In the revised manuscript, we elaborated the section based on prior studies about the development and validation of growth mindset measures and those about the relationship between growth mindset and positive youth development.

Our results from both studies suggest that the English version of the MGM measure can well measure one's MGM as we intended. In fact, the previous studies that developed and tested measurements for the mindset with diverse domains have shown that the measurements possessed good reliability and validity (e.g., Lüftenegger et al., 2015; Pomerantz & Saxon, 2001), so growth mindset can be feasible measured with self-report measures. Consistent with the previous studies about measuring growth mindset in other domains, we were able to show that the MGM can also be appropriately measured by a self-report measure, the MGM measure. Moreover, the results from our correlation analysis are consistent with findings in previous studies that have examined the positive relationship between growth mindset and

successful social adjustment and positive youth development in general (Yeager & Dweck, 2012; Yeager, Miu, Powers, & Dweck, 2013; Yeager et al., 2011). Hence, our study that tested and validated the MGM measure demonstrated that first, MGM can be well measured by the MGM measure as growth mindset in general was measured by reliable and valid tools in previous studies; and second, MGM is associated with moral and positive youth development as shown in previous growth mindset studies in other domains.

Competing Interests: I am the corresponding author of this paper.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias

- You can publish traditional articles, null/negative results, case reports, data notes and more

- The peer review process is transparent and collaborative

- Your article is indexed in PubMed after passing peer review

- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research