

**Relatable and Attainable Moral Exemplars as Sources for Moral Elevation and
Pleasantness**

Hyemin Han ^{1†} Kelsie J. Dawson ²

¹ Educational Psychology Program, University of Alabama, Alabama, AL 35487

² Maine Mathematics and Science Alliance, Augusta, ME 04330

Author ORCIDs

Hyemin Han  0000-0001-7181-2565

Kelsie J. Dawson  0000-0002-7958-3004

† **Corresponding author:** Hyemin Han, College of Education, University of Alabama, Box 870231, Tuscaloosa AL 35487, Email: hyemin.han@ua.edu

Declarations

Funding: Research reported in this publication was supported by the University of Alabama Research Grant Committee (RG14785) and by the John Templeton Foundation (48365). The content is solely the responsibility of the authors and does not necessarily represent the official views of the University of Alabama or John Templeton Foundation.

Conflicts of interest/Competing interests: The authors declare no competing financial interests.

Data availability statement: All data supporting the findings reported herein, along with the study materials and code for the statistical analyses, are available from the Open Science

Framework repository: <https://doi.org/10.17605/OSF.IO/V5NK7>

Code availability: All source codes for data analyses are available from the Open Science

Framework repository: <https://osf.io/hxuns/>

Acknowledgements: Collaboration on this research was facilitated by participation in the Summer Seminars in Neuroscience and Philosophy at Duke University, directed by Felipe De Brigard and Walter Sinnott-Armstrong. Some of this research was presented at Duke as well as the 44th annual meeting of the Association for Moral Education in Barcelona. The authors thank Joshua May, Clifford Workman, and Andrea L. Glenn for their contribution to research design and data collection.

Relatable and Attainable Moral Exemplars as Sources for Moral Elevation and Pleasantness

Abstract

In the present study, we examined how the perceived attainability and relatability of moral exemplars predicted moral elevation and pleasantness among both adult and college student participants. Data collected from two experiments were analyzed with Bayesian multilevel modeling to explore which factors significantly predicted outcome variables at the story level. The analysis results demonstrated that the main effect of perceived relatability and the interaction effect between attainability and relatability shall be included in the best prediction model, and thus, were deemed to predict the outcome variables significantly. The main effect of relatability as well as its interaction with attainability positively predicted elevation and pleasantness. We discussed educational implications of the findings in terms of how relatability may be the first point of emphasis for moral educators to focus on and attainability can then bolster the effectiveness. These relatable and attainable moral exemplars can be sources for moral elevation and pleasantness, which promote motivation to emulate moral behavior presented by the exemplars.

Keywords: moral exemplar; moral motivation; attainability; relatability; Bayesian analysis

Moral educators have regarded the stories of moral exemplars as one of the major sources for moral education (Kristjánsson, 2006). In many cases, moral exemplars are presented to students to promote their motivation to engage moral behavior through emulating presented exemplary behavior (Han et al., 2017). From the philosophical perspective, particularly that of virtue ethics, moral emulation via presenting virtuous exemplars is deemed to be a major method for moral education (Han, 2015). As exemplars are supposed to present the paragon of morals and virtues, they can provide insights about how to improve their morality with concrete examples, e.g., their actions and thoughts (Damon & Colby, 2013). In fact, exemplary stories have been widely utilized in moral education in educational settings. For instance, moral exemplars are introduced to exemplify virtues and values to be taught to students in moral education textbooks (Han, Park, Kim, et al., 2018).

Moral educators are particularly interested in how to promote moral motivation, motivation to engage in moral behavior, through educational activities. According to the traditional view of moral development focusing on moral cognition, such as the classical Kohlbergian view, moral judgment and reasoning have been regarded as the primary sources for moral motivation (Kohlberg, 1981). However, many moral educators argued that such a cognitivist view cannot well explain the gap between judgment and behavior, so additional components should be considered to address the “gappiness” issue and successfully explain the generation of moral motivation (Darnell et al., 2019). Some propose that moral identity, which is associated with whether one regards moral values as important and central to oneself, is a source of moral motivation (Hardy & Carlo, 2005). For instance, presence of moral identity has been found to promote motivation to engage in various moral and prosocial behaviors, such as donation, volunteering, civic engagement, consistently throughout one’s life (e.g., Aquino &

Reed, 2002; Han et al., 2019; Han & Dawson, 2021). Furthermore, affective aspects of moral psychology, such as empathy, also play fundamental roles in promoting moral motivation (Eisenberg & Fabes, 1990). Particularly, empathic concern, concern about others' pain and wellbeing, has been reported to promote moral motivation to help others (e.g., Decety & Yoder, 2016; Hardy, 2006). Thus, to be able to understand how to promote moral motivation effectively through moral education, it is necessary to examine the diverse sources of moral motivation.

Previous empirical studies have examined whether presenting moral exemplars was able to promote moral motivation across diverse contexts, including both experimental and educational settings. As one way to explain the mechanism based on data, Haidt (2000) mentioned moral elevation, an uplifting emotional reaction instigated by observing other's exemplary action. According to his psychological account, moral elevation plays fundamental roles in promoting one's motivation to engage in moral behavior. Follow-up experimental studies have shown that moral elevation produced by presenting exemplars resulted in promotion of prosocial motivation and behavior within diverse contexts, such as helping others without any compensation and favoring eco-friendly products (Romani et al., 2016; Schnall et al., 2010).

Although Haidt (2000) and follow-up studies have demonstrated morally exemplary stories are capable of inducing moral motivation and behavior through moral elevation, they have not provided sufficient information about which types of moral exemplars would be particularly effective. In fact, Monin (2007) argued that mere presentation of moral exemplars, particularly those seen to be extremely extraordinary so that ordinary people cannot easily emulate, is likely to backfire and produce undesirable negative outcomes, such as decrease in moral motivation and resentment. According to his point, presentation of extreme exemplars may induce negative unpleasant emotional responses, and finally, result in people withdrawing from

engaging moral behavior similar to what was presented by such exemplars (Monin et al., 2008). One possible explanation is that self-defense mechanisms may initiate once one is exposed to extreme exemplars, which are perceived to be extremely different from oneself and practically impossible to emulate (Han et al., 2017). In a recent philosophical paper, Athanassoulis (2022) also argued that *phronimos*, a perfectly virtuous person, could not be an ideal exemplar to be used in education. Instead, Athanassoulis suggested that imperfect exemplars, who seem to be accessible and attainable to students, can be more effective. These psychological and philosophical accounts suggest that specific types of moral exemplars, not all exemplars, are effective in promoting moral motivation and behavior when they are presented to ordinary people, including students.

In addition to the prior point about elevation induced by presentation of moral exemplars, we may also need to refer to works related to another dimension of emotional responses to presented exemplars, pleasantness. As Monin (2007) proposed, inappropriate introduction of extreme moral exemplars is likely to induce unpleasant emotional reactions and then decrease the level of moral motivation. In fact, works in moral philosophy, particularly those in virtue ethics, suggest that pleasantness is inseparable from moral virtues and virtuous actions. Curren and Ryan (2020) argued that to be able to induce virtuous motivation to do morally virtuous things, experiencing pleasure is essential in promoting intrinsic, not extrinsic, motives to implement moral behavior. Related to this point about the importance of pleasantness in generating virtuous motivation and behavior, within the context of moral modeling via presentation of moral exemplars, Kristjánsson (2017) suggested that admiration of moral exemplars that promotes motivation for emulation is fundamentally inseparable from a pleasant emotional response toward the presented moral exemplars. Such philosophical accounts about

pleasantness in the moral domain have been supported by empirical evidence (e.g., Jensen & Aamodt, 2002). Given these, in addition to elevation, we also need to pay attention to pleasantness as an important emotional response towards moral exemplars in explaining generation of moral motivation.

Following the above discussed psychological and philosophical frameworks, Han et al.'s (2017) study conducted in a classroom setting demonstrated that close-other exemplars, such as friends and family members, were more effective in promoting moral elevation, and finally, prosocial behavior, volunteering in particular. Their study provides useful insight about a certain type of exemplars, close-other exemplars, who are perceived to be attainable and relatable from students' perspectives, and can be more effective than others, such as historic figures and extraordinary exemplars. In a follow-up study, Han et al. (2022) conducted additional experiments to examine which factor of exemplars significantly contributed to promoting prosocial motivation. They found that perceived relatability played the most fundamental role in the motivational process. Participants who were assigned to the relatable exemplar conditions reported significantly improved prosocial motivation and behavior compared with their counterparts in the unrelatable exemplar conditions. In other words, relatable moral exemplars, who share the similar socio-cultural backgrounds with students, shall be employed in moral education for optimal outcomes. Such a significant impact of relatable models has also been reported in previous social psychological studies that addressed general motivation even out of the moral domains (e.g., Lockwood & Kunda, 1997; Rivera & Benitez, 2016).

Although the aforementioned previous works have provided researchers and educators with useful insights about how to design moral exemplar-employed moral education to maximize its motivational effect in general, they have been focusing on individual-level analyses instead of

story-level analyses. More specifically, in their studies, the unit of analysis was an individual participant, not an individual story. For instance, Han et al. (2017) and (2022) reported that participants in the close-other or relatable exemplar condition showed significantly greater motivational improvement compared with their counterparts. However, they examined changes in individuals as a function of the experimental condition, the story type, instead of which type of story was more powerful at the story level. Hence, conducting analyses focusing on individual stories' features would be necessary to identify the type of more effective exemplary stories. From the practical perspective, findings regarding which types of exemplars are effective in promoting moral motivation will be able to provide useful insights to moral educators to intend to utilize the stories of moral exemplars in moral education.

To answer the question at the story level, in the present study, we intend to examine which type of exemplary stories are more capable of inducing moral elevation and pleasant emotional reactions, which are found to predict promotion of prosocial behavioral outcomes significantly. We will conduct multilevel modeling, quantitative analysis, that enables us to examine effects at multiple different levels, to investigate the story-level effect on top of the individual-level effect (Stegmuller, 2013), which has already been tested and reported in the previous studies. This approach is appropriate to examine which types of stories are more effective while not aggregating story-level information into an individual level. The previous studies in fact analyzed individual-level data, which consisted of the mean of variables of interest (e.g., perceived elevation and pleasantness) within each individual, not the perceived elevation and pleasantness of each individual story.

Given we are primarily interested in exploring which factors significantly influence emotional outcomes, we will employ data-driven analysis based on the Bayesian perspective

(Dawson et al., 2021), not hypothesis-driven analysis, in the present study. In the majority of the prior studies examining moral exemplars, hypothesis-driven methods based on the frequentist perspective have been widely used since they were interested in comparing motivational and behavioral outcomes between different story conditions (e.g., Han et al., 2017, 2022). A significant problem that the frequentist approach in the context of data-driven analysis possesses is that results such as p -values are not capable of demonstrating whether a hypothesis of a model of interest is supported by evidence. First, such results merely show us whether a null hypothesis shall be rejected, instead of whether an alternative hypothesis of interest shall be accepted (Wagenmakers et al., 2018). Second, in the case of ordinary frequentist regression analysis, only one model is being tested, so the result does not give us any assurance about whether the model is the best among all possible models generated from collected data (Han, 2022b). On the other hand, Bayesian analysis can suggest the extent to which a specific hypothesis or model is supported by evidence, and the model is better than its alternatives (Dawson et al., 2021).

Unlike the majority of the previous studies in the field, we intend to examine the best model predicting elevation and pleasantness with data, so conventional frequentist inference could not be an ideal approach. Hence, we will use Bayesian multilevel modeling, which allows us to compare possible prediction models and identify the best model given data (Rachev et al., 2021). Through the data-driven exploration process, we will examine which factors, perceived attainability and relatability, and the interaction effect between them, significantly predict emulation and pleasantness at the story level.

Methods

Datasets

We analyzed the datasets collected by Han et al. (2022) (see Han et al. (2022) for further details about the data descriptions and experimental procedures including presented exemplary stories). The dataset is available to the public following open science guidelines via the Open Science Framework at <https://doi.org/10.17605/OSF.IO/V5NK7>. In their previous study, they examined whether participants' behavioral outcomes were significantly altered after being presented with exemplary stories across different experimental conditions. The datasets consist of two subsets: one collected from Amazon mTurkers and one from college students at a public university located in the Southern United States. The first dataset, the mTurk dataset, includes responses from 401 participants (44.64% female; mean age = 34.99 years, SD = 10.17 years). The second dataset, the college student dataset, was collected from 218 participants (87.16% female; mean age = 20.49 years, SD = 6.67 years).

In the previous study, the participants were randomly assigned one of five conditions: attainable/relatable, attainable/non-relatable, unattainable/relatable, unattainable/non-relatable, and non-moral conditionsⁱ. According to their condition assignment, they were presented with different types of 26 exemplary stories (e.g., (un)attainable/(non-)relatable moral stories and non-moral stories). These stories were originally obtained from a previous study (Knutson et al., 2010), and then modified to manipulate attainability and relatability. In the present study, responses collected from participants assigned to the non-moral story group were excluded as the reported elevation and pleasantness from the group were not relevant to the moral domain; the group was excluded from analyses following Han et al. (2022). After presenting each story, a set of four questions, the perceived attainability and relatability of the presented exemplar(s), and

the degree of evoked elevation and pleasantness, were asked. Participants' responses were anchored to a seven-point Likert scale. These are the four items:

Perceived attainability: How difficult do you think it would be to do the same things as the person described in the story? (1: Not difficult at all – 7: Extremely difficult; reverse coded)

Perceived relatability: How similar do you think your cultural and social background is to the person described in the story? (1: Not at all similar– 7: Extremely similar)

Moral elevation: The story made me feel morally elevated (warm, uplifted - like when seeing unexpected acts of human goodness, kindness, or compassion). (1: Strongly disagree – 7: Strongly agree)

Pleasantness: How pleasant do you find the actions of the person described in the story? (1: Extremely unpleasant – 7: Extremely pleasant)

Statistical Analysis

In the present study, we utilized R for intended statistical analyses. For replicability, all source code and data files are available via the Open Science Framework at <https://osf.io/hxuns/>.

Descriptive Statistics and Correlational Analysis

For basic information about the dataset, we examined descriptive statistics of the datasets, the range, mean, median, and standard deviation of variables of interest, i.e., perceived attainability, relatability, moral elevation, and pleasantness. Furthermore, to examine how each variable was associated with each other, correlation analysis was also performed. The aforementioned information was acquired for the whole data as well as for each individual dataset.

Bayesian Multilevel Modeling

Bayesian multilevel modeling was conducted with an R package, *brms* (Bürkner, 2017). Given participants' responses were nested in multiple levels, i.e., the dataset, experimental group, and individual participant levels, since each participant was presented with 26 stories, we included the random effects of the dataset number (dataset 1 or 2), group assignment (one of four experimental groups), and participant IDs in the analyzed models. In a simultaneous manner, we examined predictors of interest, the perceived attainability and relatability of each story, in the models. For model exploration, we also tested the interaction effect between the aforementioned two predictors as well.

In the present study, we used Bayes Factors (BFs) of prediction models to identify the best model predicting outcome variables of interest, moral elevation and pleasantness (Kass & Raftery, 1995). A BF indicates to what extent one model is more strongly supported by evidence compared with another model (Han, Park, & Thoma, 2018; Wagenmakers et al., 2018). For instance, BF_{AB} demonstrates how Model A is more strongly supported by evidence compared with Model B. While interpreting BFs, we examined $\log(BF)$. According to Han et al. (2018), $2\log(BF) \geq 2$ suggests presence of positive evidence, $2\log(BF) \geq 6$ strong evidence, and $2\log(BF) \geq 10$ very strong evidence. BFs of different models were compared to identify which model best predicted outcome variables at the story level.

In the present study, we compared three different models, i.e., M0: the null model only with random effects; M1: the partial model with two main effects of attainability and relatability; M2: the full model with all main and interaction effects. In terms of *brms* grammar, M0, M1, and M2 were defined as follows:

$$M0: DV \sim (1|dataset/group/ID)$$

$$M1: DV \sim \text{attainability} + \text{relatability} + (1|\text{dataset/group/ID})$$

$$M2: DV \sim \text{attainability} + \text{relatability} + \text{attainability} \times \text{relatability} + (1|\text{dataset/group/ID})$$

where DV is a dependent variable being tested, moral elevation or pleasantness, $dataset$ is a dataset number, $group$ is an experiment group assignment, ID is an individual participant's ID. Furthermore, to examine whether random slopes of the predictors in addition to the random intercepts, we also tested the following model as well:

$$M3: DV \sim \text{attainability} + \text{relatability} + \text{attainability} \times \text{relatability} + (1+\text{attainability}+\text{relatability}|\text{dataset/group/ID})$$

To identify the best model among these candidate models, we calculated BF_{10} , BF_{20} , and BF_{30} , which indicated BF of M_1 vs. M_0 , BF of M_2 vs. M_0 , and BF of M_3 vs. M_0 , respectively (Han, 2022a). Then, for interpretation, we converted those BF values into $2\log(\text{BF})$ s. These three $2\log(\text{BF})$ s were used to compare the null model with the three other models with effects of interest. For instance, M_1 and M_2 were compared by calculating $BF_{21} = BF_{20} / BF_{10}$. Because we used $2\log(\text{BF})$ values for interpretation, two models can be compared as follows: $2\log(\text{BF}_{21}) = 2\log(\text{BF}_{20} / \text{BF}_{10}) = 2\log(\text{BF}_{20}) - 2\log(\text{BF}_{10})$. In the similar manner, we also calculated $2\log(\text{BF}_{31})$ and $2\log(\text{BF}_{32})$ as well. We examined whether calculated $2\log(\text{BF})$ s exceeded at least 2, the threshold for positive evidence.

Once the best model was identified by comparing model BFs, we examined whether each tested effect, attainability, relatability, and interaction between these two, was significantly greater than zero in the identified best model. For this purpose, we calculated and tested BF_{10} , which indicated to what extent an alternative hypothesis, the effect of interest is greater than zero, was more strongly supported by evidence than a null hypothesis. Similar to the case of the

interpretation of model BFs, $BF \geq 3$, 10, and 100 were used for thresholds for positive, strong, and very strong evidence, respectively.

Results

Descriptive Statistics and Correlational Analysis

Descriptive statistics of attainability, relatability, elevation, and pleasantness are presented in Table 1.

Table 1

Descriptive statistics of analyzed variables

Dataset	Attainability				Relatability				Elevation				Pleasantness			
	Range	M	Median	SD	Range	M	Median	SD	Range	M	Median	SD	Range	M	Median	SD
All	[1 7]	4.03	4.00	1.97	[1 7]	3.80	4.00	1.75	[1 7]	5.22	5.00	1.59	[1 7]	5.53	6.00	1.40
mTurk	[1 7]	3.94	4.00	2.00	[1 7]	3.86	4.00	1.75	[1 7]	5.19	5.00	1.59	[1 7]	5.59	6.00	1.32
College student	[1 7]	4.18	4.00	1.92	[1 7]	3.72	4.00	1.75	[1 7]	5.26	5.00	1.59	[1 7]	5.45	6.00	1.51

The result of correlation analysis is demonstrated in Table 2.

Table 2

Correlation between analyzed variables

	Attainability	Relatability	Elevation
Relatability	.17		
Elevation	-.02*	.27***	
Pleasantness	.01	.24***	.74***

Note. *: $p < .05$. ***: $p < .001$. False positive discovery rate correction applied.

In the conducted correlation test, because each individual story was the unit of the analysis, the total number of examined data points was very large ($> 10,000$). Thus, even a small effect resulted in a p -value smaller than .05, so the result shall be interpreted with caution (Han,

Park, & Thoma, 2018). Although the p -value of the association between attainability and elevation was smaller than .05, the correlation coefficient was very small, .02. It may indicate that the correlation was not practically meaningful at all. Instead, reliability, elevation, and pleasantness were significantly associated with each other, given the resultant correlation coefficients were greater than .20, which is a threshold for a small effect.

Bayesian Multilevel Modeling

Elevation

When elevation was analyzed as a dependent variable, the result of Bayesian multilevel modeling indicated that the full model with all main and interaction effects and random slopes (M3) was the best model compared with the null model (M0), partial model (M1), and random intercept-only full model (M2).

First, when the null model was compared with the three other models, $2\log(\text{BF}_{10}) = 540.42$, $2\log(\text{BF}_{20}) = 662.41$, and $2\log(\text{BF}_{30}) = \infty$. When M1, M2, and M3 were compared with each other, $2\log(\text{BF}_{21}) = 121.98$, $2\log(\text{BF}_{31}) = \infty$, and $2\log(\text{BF}_{32}) = \infty$. Given $2\log(\text{BF}_{30})$, $2\log(\text{BF}_{31})$, and $2\log(\text{BF}_{32})$ were extremely large, M3, the full model including both random slopes and intercepts was found to be best supported by evidence.

Table 3

Coefficients of interest estimated by Bayesian multilevel modeling with the full model including all random intercepts and slopes (M3).

	Estimated value	Error	95% credible interval	
			Lower bound	Upper bound
Dependent variable: Elevation				
Attainability	-.05	.23	-.55	.42

Relatability	.22	.13	-.05	.47
Attainability x Relatability	.03	.01	.02	.05
Dependent variable: Pleasantness				
Attainability	-.01	.17	-.38	.33
Relatability	.20	.18	-.19	.51
Attainability x Relatability	.05	.01	.03	.07

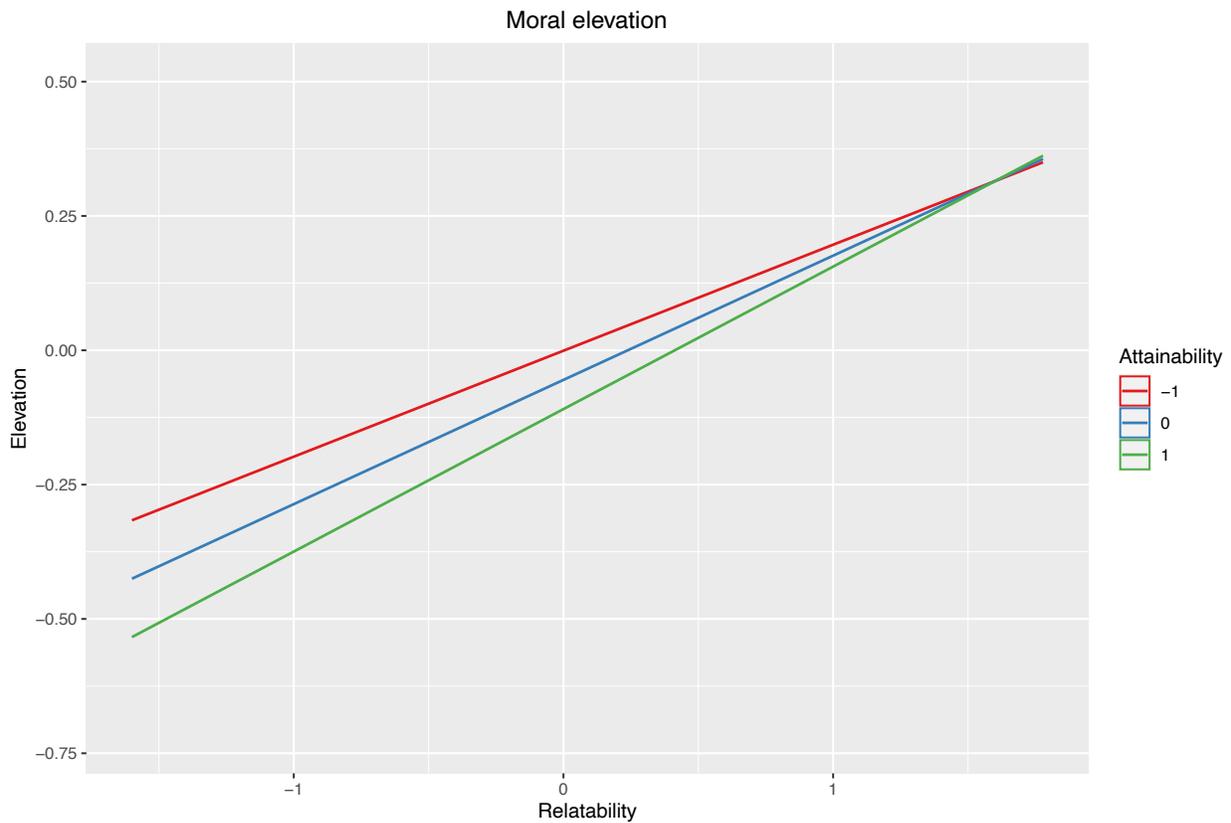
Second, we also examined whether the main effects of attainability and relatability, and the interaction effect between them were significantly greater than zero in M3 with BFs (see Table 3 for estimated coefficients of interest). When attainability was examined, the resultant $2\log(\text{BF})$ was -1.55. It suggests that evidence was not sufficient to support the alternative hypothesis that the main effect of attainability was significantly greater than zero. Additional exploratory analysis was conducted to examine whether the effect of attainability was significantly different from zero. The resultant $2\log(\text{BF}) = -1.01$ indicating that the main effect of attainability was not significantly different from zero. On the other hand, the main effect of relatability was deemed to be greater than zero given $2\log(\text{BF}) = 6.78$ indicating presence of strong evidence. The higher relatability was associated with higher elevation. Presence of the significant positive interaction effect was very strongly supported by evidence, $2\log(\text{BF}) = 18.42$.

Figure 1 demonstrates the result of simple slope analysis, which was employed to visualize the interaction effect within the context of multilevel modeling (Preacher et al., 2006). The direction of the interaction suggested that the attainability and relatability supported each

other's positive association with elevation. The elevation outcome was maximized when a presented story was perceived to be attainable as well as relatable.

Figure 1

Interaction between relatability and attainability predicting moral elevation.



Note. All variables, i.e., relatability, attainability, elevation, were standardized. Red, blue, and green lines with three different slopes represent the association between relatability and moral elevation when attainability was -1 standard deviation, 0 standard deviation, and +1 standard deviation, respectively.

Pleasantness

We also examined the candidate prediction models for pleasantness as well. First, when the four models were tested, similar to the case of attainability, the full model including both random slopes and intercepts (M3) was found to be best supported by evidence. When M1, M2,

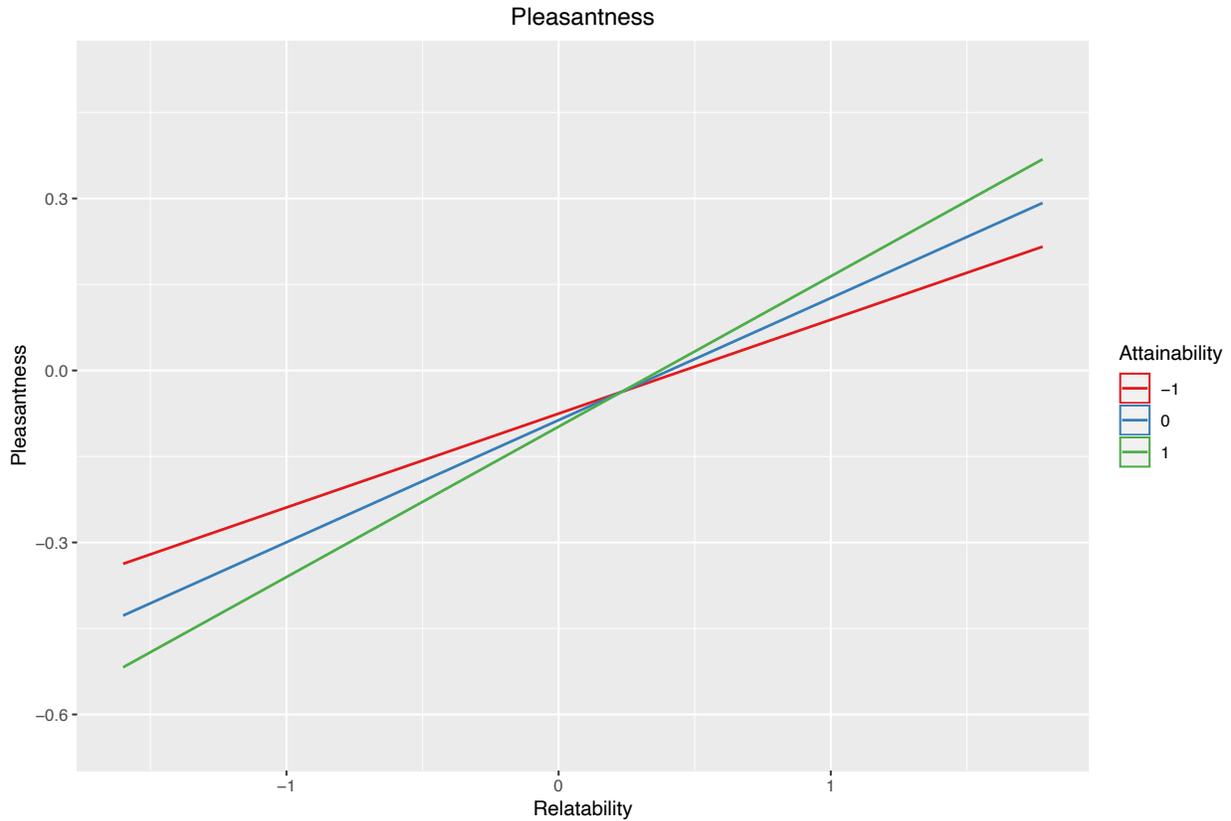
and M3 were compared with M0, the resultant $2\log(\text{BF}_{10}) = 449.96$, $2\log(\text{BF}_{20}) = 572.68$, and $2\log(\text{BF}_{30}) = 1302.62$. As a result, $2\log(\text{BF}_{21}) = 122.71$, $\log(\text{BF}_{31}) = 852.66$, and $2\log(\text{BF}_{32}) = 729.95$. Because all $2\log(\text{BF}_{30})$, $2\log(\text{BF}_{31})$, and $2\log(\text{BF}_{32})$ sufficiently exceeded 10, a thresholding for presence of very strong evidence, we shall conclude that M3 was the best model predicting pleasantness among the candidate models.

Second, we tested the main and interaction effects in M3 as well (see Table 3 for estimated coefficients of interest). When the main effect of attainability was examined, the resultant $2\log(\text{BF})$ was $-.52$ indicating that evidence was insufficient to support the alternative hypothesis. As additional exploratory analysis, we also examined whether the main effect of attainability was significantly different from zero. The calculated $2\log(\text{BF}) = -3.62$ suggesting that there was evidence positively supporting that the effect of attainability was zero. On the other hand, the main effect of relatability was deemed to be greater from zero because the resultant $2\log(\text{BF}) = 5.34$, so evidence positively supported the alternative hypothesis. The higher relatability predicted the higher pleasantness. Presence of the significant positive interaction effect was very strongly supported by evidence given $2\log\text{BF}$ was ∞ .

Figure 2 shows the result of simple slope analysis. Identical to the case of elevation, there was a significant positive interaction between attainability and relatability. The highest pleasant outcome resulted when both perceived attainability and relevance were high.

Figure 2

Interaction between relatability and attainability predicting pleasantness.



Note. All variables, i.e., reliability, attainability, elevation, were standardized. Red, blue, and green lines with three different slopes represent the association between reliability and pleasantness when attainability was -1 standard deviation, 0 standard deviation, and +1 standard deviation, respectively.

Exploratory Analysisⁱⁱ

While examining the correlation between the variables of interest, we discovered two interesting patterns: 1. The correlation between the two predictors, perceived attainability and reliability, was weak, $r = .17$; and 2. The correlation between the two outcome variables, elevation and pleasantness, was strong, $r = .74$. We assume that the first interesting pattern, the lack of association between the two predictors, might be attributable to the nature of the original experiments and presented materials. When Han et al. (2022) prepared exemplary stories according to different attainability and reliability group assignments, they intentionally

manipulated such aspects of the presented moral exemplars. In fact, Han et al.'s (2022) Study 2A, a norming study, demonstrated that participants reported significantly different perceived attainability and relatability according to each story's group assignment as intended. Given this, it would be possible to expect that the perceived attainability and relatability in the present study was also significantly differentiated, so the correlation between them was weak accordingly.

Given the two outcome variables were very strongly associated with each other, we examined whether the result of the multilevel modeling would be significantly altered if the two outcome variables were combined into one. To examine this possibility, we first calculated the sum score of elevation plus pleasantness. Then, we conducted the same Bayesian multilevel modeling while considering the sum score as the dependent variable. Among the tested models, similar to the prior analysis, the model with random intercepts and slopes was reported to be the best, $2\log(\text{BF}_{10}) = 696.35$, $2\log(\text{BF}_{20}) = 719.79$, and $2\log(\text{BF}_{30}) = 1759.43$. The main effect of relatability was significant, $b = .41$, $SE = .25$, 95% Bayesian CI [.06, .70], $2\log\text{BF} = 6.37$, while that of attainability was non-significant, $b = -.06$, $SE = .29$, 95% Bayesian CI [-.47, .39], $2\log\text{BF} = -1.27$. The interaction effect was significant, $b = .08$, $SE = .02$, 95% Bayesian CI [.05, .11], $2\log\text{BF} = \infty$. Both the main effect of relatability and interaction effect were positive identical to the results from the main analyses.

Discussion

In the present study, we explored the best models predicting moral elevation and pleasantness evoked by watching different types of moral exemplary stories. The model exploration was conducted at the story level, not the individual level, with Bayesian multilevel modeling to focus on the story-level factors, perceived attainability and relatability since such factors were our primary interests. Unlike the previous studies examining the effectiveness of

different types of moral exemplary stories, i.e., Han et al. (2017) and (2022), which concentrated upon individual-level differences by comparing outcome across different experimental conditions, we were able to test the effect of each story's attainability and relatability. Furthermore, we also examined one additional candidate prediction model including the interaction effect between attainability and relatability, which has not been yet tested in the previous studies.

The conducted Bayesian multilevel modeling identified two best prediction models, one for moral elevation and one for pleasantness, successfully. The calculated model BFs indicated that for both dependent variables, the full model with all main and interaction effects as well as random intercepts and slopes was best supported by evidence. When each predictor was examined, in both models, the main effect of relatability and the interaction effect between attainability and relatability were found to be significantly greater than zero. Both the main effect of relatability and the interaction effect positively predicted moral elevation and pleasantness. The main effect of attainability was not significantly greater than zero. Following additional exploratory analyses suggested that the effect of attainability was in fact not significantly different from zero.

In general, the findings were consistent with Han et al. (2022), which reported the significant association between perceived relatability, and emotional and motivational outcomes at the individual level. Han et al. (2022) also found that attainability was non-significant in predicting the outcomes. Perhaps, the used stories presented different types of demands and efforts to engage in moral behavior (e.g., money, time), so it would be difficult to manipulate perceived attainability accurately as intended. This methodological limitation involving attainability manipulation might result in the non-significance of perceived attainability in the

experiments. On the other hand, relatability was relatively easier and more straightforward to manipulate. Given all participants were Americans, it was possible to manipulate perceived relatability by altering the presented exemplars' nationality and cultural backgrounds unlike the case of attainability manipulation. Hence, the main effect of relatability might be the only significant main effect in predicting outcome variables.

Although the finding was consistent with Han et al. (2022) when main effects were considered, we also found that the interaction effect between attainability and relatability was also significant in the identified models. As shown by the simple slope analysis results, the positive impact of relatable exemplars became stronger when such exemplars were perceived to be attainable. Such an effect was demonstrated by the steeper slope when perceived attainability was high (+1 standard deviation). Given the previous study was not able to examine the interaction effect as it was primarily interested in comparing outcomes across conditions, it would be a novel finding from the present study. In general, this significant interaction effect may suggest that attainability works as a booster to strengthen the positive association between relatability and positive emotional outcomes while it could not become the significant independent predictor.

One point that we may need to note is that attainability itself as a main effect could not significantly contribute to predicting elevation and pleasantness; instead, it was able to predict the outcomes only through interacting with relatability. As one possible theoretical explanation of the result, in addition to the aforementioned methodological limitation related to the difficulty to manipulate attainability, social comparison proposed in social psychology can be considered. According to previous research, social comparison occurs when one is comparing their abilities with others within social contexts (Festinger, 1954). The result of social comparison influences

one's motivation (Suls et al., 2002). In general, when one is being compared with superior others, the perceived gap is likely to promote motivation for self-improvement to fill the perceived gap (Blanton et al., 1999; Huguet et al., 2001). Han et al. (2017, 2022) also suggested that social comparison may play an important role in motivation generation through presenting moral exemplars. Such social comparison occurs more vigorously when one is comparing oneself with others who are perceived to be similar to them (Tsay-Vogel & Krakowiak, 2019). For instance, when the target of social comparison is deemed to share similar backgrounds, such as gender and ethnicity, with oneself, then the emotional and motivational impact of such comparison becomes significantly stronger than when the target is not deemed to be similar (Garcia et al., 2013). Hence, perceived relatability as a main effect would be the most significant predictor of emotional outcomes as whether the presented exemplar is perceived to be similar to oneself determines the degree of social comparison, and finally, that of psychological impacts before attainability becomes a factor to be considered.

The significant interaction effect between attainability and relatability can be explained by the aforementioned mechanism of social comparison (Lockwood & Kunda, 1997). Once an exemplar is perceived to be relatable, then emotional and motivational outcomes get significantly influenced by attainability as a result of such comparison (Zanna et al., 1975). In this situation, exemplars who are perceived to be attainable are more likely to generate positive emotional responses, elevation and pleasantness as found in the present study. This result is consistent with what has been reported in previous studies, such as Han et al. (2017) that directly compared attainable and unattainable exemplars while controlling for perceived relatability (see Experiment 1 in Han et al. (2017)). Because Han et al.'s (2017) Experiment 1 only compared relatable attainable versus unattainable exemplars, they were only able to examine the effect of

attainability while perceived relatability commonly existed across conditions. The finding from the present study may provide additional information regarding the point. As mentioned, if attainability would become a factor positively contributing to motivational promotion, then relatability is required as a prerequisite within the context of exemplar presentation. Perhaps, Han et al. (2017) was able to show the influence of attainability because all the presented exemplars, relatable exemplars, sufficed the prerequisite.

The findings from the present study may provide practical implications for moral education utilizing moral exemplars. Consistent with what has been reported in Han et al. (2017, 2022), in the first place, moral educators need to consider employing relatable exemplars, those who are sharing the similar socio-cultural backgrounds with students. Even if exemplars present doable exemplary behaviors, their motivational effects may be significantly limited when they are unrelatable from students' perspectives. After assuring perceived relatability of exemplars, then, it would become important to examine whether presented exemplary behaviors are deemed to be emulatable with reasonable efforts. As demonstrated in Figures 1 and 2 in the present study, educators shall consider relatable as well as attainable exemplars to maximize the positive emotional and motivational outcomes within the context of moral education.

Although findings from the present study may provide useful insights for future research as well as improvement of moral education, several limitations warrant additional studies. First, as mentioned earlier, manipulation of attainability could be a significant methodological limitation. Because different situations were presented across 26 stories, it was difficult to manipulate attainability of each exemplary behavior in a consistent manner. Second, unlike previous studies, such as Han et al. (2017, 2022), we were not able to examine behavioral outcomes, such as volunteering and donation, as dependent variables. Because such behavioral

outcomes were measured at the individual level, not at the story level, they could not be analyzed in the present study, which employed multilevel modeling.

Concluding Remarks

The attainability and relatability of the presented moral exemplars have been regarded as factors determining the effectiveness of such exemplars in promoting moral motivation among students. In general, to be able to maximize their effectiveness, moral exemplary stories should be attainable and relatable from students' perspectives. Empirical evidence presented in the present study may provide further insights about how attainability and relatability work at the story level. As demonstrated, relatability can be considered as a necessary condition for effectively motivating students via presentation of moral exemplars. Furthermore, attainability was found to be capable of boosting the effect of relatability in the process. Thus, moral educators may need to take into account the above-mentioned mechanism of relatability and attainability while implementing exemplars in moral education.

References

- Aquino, K., & Reed, A. (2002). The self-importance of moral identity. *Journal of Personality and Social Psychology*, 83(6), 1423–1440. <https://doi.org/10.1037/0022-3514.83.6.1423>
- Athanassoulis, N. (2022). The Phronimos as a moral exemplar: Two internal objections and a proposed solution. *The Journal of Value Inquiry*. <https://doi.org/10.1007/s10790-021-09872-4>
- Blanton, H., Buunk, B. P., Gibbons, F. X., & Kuyper, H. (1999). When better-than-others compare upward: Choice of comparison and comparative evaluation as independent predictors of academic performance. *Journal of Personality and Social Psychology*, 76(3), 420–430. <https://doi.org/10.1037/0022-3514.76.3.420>

- Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*(1). <https://doi.org/10.18637/jss.v080.i01>
- Curren, R., & Ryan, R. M. (2020). Moral self-determination: The nature, existence, and formation of moral motivation. *Journal of Moral Education*, *49*(3), 295–315. <https://doi.org/10.1080/03057240.2020.1793744>
- Damon, W., & Colby, A. (2013). Why a true account of human development requires exemplar research. In M. K. Matsuba, P. E. King, & K. C. Bronk (Eds.), *Exemplar methods and research: Quantitative and qualitative strategies for investigation. New Directions in Child and Adolescent Development* (pp. 13–26). Jossey-Bass.
- Darnell, C., Gulliford, L., Kristjánsson, K., & Paris, P. (2019). Phronesis and the Knowledge-Action Gap in Moral Psychology and Moral Education: A New Synthesis? *Human Development*, *62*(3), 101–129. <https://doi.org/10.1159/000496136>
- Dawson, K. J., Han, H., & Choi, Y. R. (2021). How are moral foundations associated with empathic traits and moral identity? *Current Psychology*. <https://doi.org/10.1007/s12144-021-02372-5>
- Decety, J., & Yoder, K. J. (2016). Empathy and motivation for justice: Cognitive empathy and concern, but not emotional empathy, predict sensitivity to injustice for others. *Social Neuroscience*, *11*(1), 1–14. <https://doi.org/10.1080/17470919.2015.1029593>
- Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, *14*(2), 131–149. <https://doi.org/10.1007/BF00991640>
- Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*, 117–140. <https://doi.org/10.1177/001872675400700202>

- Garcia, S. M., Tor, A., & Schiff, T. M. (2013). The Psychology of Competition: A Social Comparison Perspective. *Perspectives on Psychological Science*, 8(6), 634–650.
<https://doi.org/10.1177/1745691613504114>
- Haidt, J. (2000). The positive emotion of elevation. *Prevention and Treatment*, 3(3), 1–5.
- Han, H. (2015). Virtue ethics, positive psychology, and a new model of science and engineering ethics education. *Science and Engineering Ethics*, 21(2), 441–460.
<https://doi.org/10.1007/s11948-014-9539-7>
- Han, H. (2022a). Trust in the scientific research community predicts intent to comply with COVID-19 prevention measures: An analysis of a large-scale international survey dataset. *Epidemiology and Infection*, 150, e36.
<https://doi.org/10.1017/S0950268822000255>
- Han, H. (2022b). Exploring the relationship between purpose and moral psychological indicators. *Ethics & Behavior*, 1–12. <https://doi.org/10.1080/10508422.2022.2142227>
- Han, H., & Dawson, K. J. (2021). Applying elastic-net regression to identify the best models predicting changes in civic purpose during the emerging adulthood. *Journal of Adolescence*, 93, 20–27. <https://doi.org/10.1016/j.adolescence.2021.09.011>
- Han, H., Kim, J., Jeong, C., & Cohen, G. L. (2017). Attainable and Relevant Moral Exemplars Are More Effective than Extraordinary Exemplars in Promoting Voluntary Service Engagement. *Frontiers in Psychology*, 8, 283. <https://doi.org/10.3389/fpsyg.2017.00283>
- Han, H., Liauw, I., & Kuntz, A. F. (2019). Moral Identity Predicts the Development of Presence of Meaning During Emerging Adulthood. *Emerging Adulthood*, 7(3), 230–237.
<https://doi.org/10.1177/2167696818758735>

- Han, H., Park, J., & Thoma, S. J. (2018). Why do we need to employ Bayesian statistics and how can we employ it in studies of moral education?: With practical guidelines to use JASP for educators and researchers. *Journal of Moral Education*, 47(4), 519–537.
<https://doi.org/10.1080/03057240.2018.1463204>
- Han, H., Park, S. C., Kim, J., Jeong, C., Kunii, Y., & Kim, S. (2018). A quantitative analysis of moral exemplars presented in moral education textbooks in Korea and Japan. *Asia Pacific Journal of Education*. <https://doi.org/10.1080/02188791.2018.1423950>
- Han, H., Workman, C. I., May, J., Scholtens, P., Dawson, K. J., Glenn, A. L., & Meindl, P. (2022). Which moral exemplars inspire prosociality? *Philosophical Psychology*, 1–28.
<https://doi.org/10.1080/09515089.2022.2035343>
- Hardy, S. A. (2006). Identity, reasoning, and emotion: An empirical comparison of three sources of moral motivation. *Motivation and Emotion*, 30(3), 207–215.
- Hardy, S. A., & Carlo, G. (2005). Identity as a source of moral motivation. *Human Development*, 48(4), 232–256.
- Huguet, P., Dumas, F., Monteil, J. M., & Genestoux, N. (2001). Social comparison choices in the classroom: Further evidence for students' upward comparison tendency and its beneficial impact on performance. *European Journal of Social Psychology*, 31(5), 557–578.
<https://doi.org/10.1002/ejsp.81>
- Jensen, K., & Aamodt, P. O. (2002). Moral motivation and the battle for students: The case of studies in nursing and social work in Norway. *Higher Education*, 44(3/4), 361–378.
<https://doi.org/10.1023/A:1019817601131>
- Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795. <https://doi.org/10.2307/2291091>

- Knutson, K. M., Krueger, F., Koenigs, M., Hawley, A., Escobedo, J. R., Vasudeva, V., Adolphs, R., & Grafman, J. (2010). Behavioral norms for condensed moral vignettes. *Social Cognitive and Affective Neuroscience, 5*(4), 378–384.
<https://doi.org/10.1093/scan/nsq005>
- Kohlberg, L. (1981). *The philosophy of moral development: Moral stages and the idea of justice*. Harper & Row.
- Kristjánsson, K. (2006). Emulation and the use of role models in moral education. *Journal of Moral Education, 35*(1), 37–49.
- Kristjánsson, K. (2017). Emotions targeting moral exemplarity: Making sense of the logical geography of admiration, emulation and elevation. *Theory and Research in Education, 15*(1), 20–37. <https://doi.org/10.1177/1477878517695679>
- Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology, 73*, 91–103.
<https://doi.org/10.1037/0022-3514.73.1.91>
- Monin, B. (2007). Holier than me? Threatening social comparison in the moral domain. *Revue Internationale de Psychologie Sociale, 20*, 53–68.
- Monin, B., Sawyer, P. J., & Marquez, M. J. (2008). The rejection of moral rebels: Resenting those who do the right thing. *Journal of Personality and Social Psychology, 95*, 76–93.
<https://doi.org/10.1037/0022-3514.95.1.76>
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational Tools for Probing Interactions in Multiple Linear Regression, Multilevel Modeling, and Latent Curve Analysis. *Journal of Educational and Behavioral Statistics, 31*(4), 437–448.
<https://doi.org/10.3102/10769986031004437>

- Rachev, N. R., Han, H., Lacko, D., Gelpí, R., Yamada, Y., & Lieberoth, A. (2021). Replicating the Disease framing problem during the 2020 COVID-19 pandemic: A study of stress, worry, trust, and choice under risk. *PLOS ONE*, *16*(9).
<https://doi.org/10.1371/journal.pone.0257151>
- Rivera, L. M., & Benitez, S. (2016). The Roles of In-Group Exemplars and Ethnicracial Identification in Self-Stereotyping. *Social Cognition*, *34*(6), 604–623.
<https://doi.org/10.1521/soco.2016.34.6.604>
- Romani, S., Grappi, S., & Bagozzi, R. P. (2016). Corporate Socially Responsible Initiatives and Their Effects on Consumption of Green Products. *Journal of Business Ethics*, *135*(2), 253–264. <https://doi.org/10.1007/s10551-014-2485-0>
- Schnall, S., Roper, J., & Fessler, D. M. T. (2010). Elevation leads to altruistic behavior. *Psychological Science*, *21*, 315–320. <https://doi.org/10.1177/0956797609359882>
- Stegmueller, D. (2013). How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches. *American Journal of Political Science*, *57*(3), 748–761. <https://doi.org/10.1111/ajps.12001>
- Suls, J., Martin, R., & Wheeler, L. (2002). Social comparison: Why, with whom, and with what effect? *Current Directions in Psychological Science*, *11*, 159–163.
<https://doi.org/10.1111/1467-8721.00191>
- Tsay-Vogel, M., & Krakowiak, K. M. (2019). The virtues and vices of social comparisons: Examining assimilative and contrastive emotional reactions to characters in a narrative. *Motivation and Emotion*, *43*(4), 636–647. <https://doi.org/10.1007/s11031-019-09756-y>
- Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Selker, R., Gronau, Q. F., Dropmann, D., Boutin, B., Meerhoff, F., Knight, P., Raj, A., van Kesteren, E.-J.,

van Doorn, J., Šmíra, M., Epskamp, S., Etz, A., Matzke, D., ... Morey, R. D. (2018).

Bayesian inference for psychology. Part II: Example applications with JASP.

Psychonomic Bulletin & Review, 25(1), 58–76. [https://doi.org/10.3758/s13423-017-1323-](https://doi.org/10.3758/s13423-017-1323-7)

7

Zanna, M. P., Goethals, G. R., & Hill, J. F. (1975). Evaluating a sex-related ability: Social

comparison with similar others and standard setters. *Journal of Experimental Social*

Psychology, 11(1), 86–93. [https://doi.org/10.1016/S0022-1031\(75\)80013-8](https://doi.org/10.1016/S0022-1031(75)80013-8)

ⁱ In the non-moral condition, Han et al. (2022) presented the stories of exemplars that may inspire participants but not in a moral way (e.g., a professional car racer who won competitions even under very threatening conditions, a violinist who became a successful artist after diligent exercises, a passionate student journalist who successfully published books and articles, etc.). These stories were designed to present excellences in diverse domains rather than moral virtues.

ⁱⁱ The exploratory analysis was conducted according to an anonymous reviewer's comments and suggestions on the earlier version of the manuscript. We appreciate their invaluable feedback to improve the quality of our work.