JOURNAL OF RESEARCH METHODOLOGY

Journal of Research Methodology
Vol. 37 No. 1 (January–April 2024)
ISSN (online) 2697-4835
Research Article https://doi.org/10.14456/jrm.2024.1

Analysis of Growth Curve Modeling of Students' Moral Reasoning on Applications of Modern Biotechnology

Onrumpa Kumnuanek[1], Pongprapan Pongsopon[2], and Uriwan Aranyawat[3]

Received: July 15, 2022 – Revised: March 6, 2023 – Accepted: March 9, 2023

Abstract

This study examined the growth of students' moral reasoning on the moral dilemma of applications of modern biotechnology. A total of 206 high school students participated in this study. They underwent four waves of assessment over one semester. We validated a hypothesized model of longitudinal data using a multilevel analysis framework by Mplus 8.0. The results showed that 32.6% of the total variance was due to individual differences (intraclass coefficient = .326). There was variation in terms of students' initial moral reasoning scores ($\sigma^2_{\mu 0}$ = 2.719, p < .001). At an intrapersonal level, moral reasoning can be explained by time and moral sensitivity. Moral reasoning would increase by 0.19 point in each subsequent measurement (p < .05). At an interpersonal level, the mean of initial moral reasoning was 11.33 points (p < .01) and at the initial point, girls scored 1.377 points more than boys on the moral reasoning measure. Knowledge of DNA technologies and previous biology achievement could not explain the within- and between-person variance of moral reasoning, respectively. The pedagogical implications for moral education were also discussed.

Keywords: Application of Modern Biotechnology, Growth Curve Modeling, Moral Reasoning

[1] Department of Education, Faculty of Education, Kasetsart University, Bangkok, 10900, Thailand.
onrumpa.ku@ku.th

[2] *Corresponding Author,*
Department of Education, Faculty of Education, Kasetsart University, Bangkok, 10900, Thailand.
pongprapan.p@ku.th

[3] Department of Genetics, Faculty of Science, Kasetsart University, Bangkok, 10900, Thailand.
fsciuwa@ku.ac.th

วารสารวิธีวิทยาการวิจัย
ปีที่ ๓๗ ฉบับที่ ๑ (มกราคม–เมษายน ๒๕๖๗)
ISSN (online) 2697-4835
https://doi.org/10.14456/jrm.2024.1 บทความวิจัย

การวิเคราะห์โมเดลโค้งพัฒนาการการให้เหตุผลเชิงจริยธรรมของนักเรียนที่มีต่อการประยุกต์ใช้เทคโนโลยีชีวภาพสมัยใหม่

อรรัมภา คำนูเอนก[1], พงศ์ประพันธ์ พงษ์โสภณ[2], และ อุไรวรรณ อรัญวาสน์[3]

รับต้นฉบับ : 15 กรกฎาคม 2565 – รับแก้ไข : 6 มีนาคม 2566 – ตอบรับตีพิมพ์ : 9 มีนาคม 2566

บทคัดย่อ

งานวิจัยนี้ศึกษาการเติบโตของการให้เหตุผลเชิงคุณธรรมในนักเรียนเกี่ยวกับประเด็นการประยุกต์ใช้เทคโนโลยีชีวภาพสมัยใหม่ นักเรียนระดับมัธยมปลายจำนวน 206 คนเข้าร่วมในงานวิจัยนี้ นักเรียนถูกวัดการให้เหตุผลเชิงคุณธรรมจำนวน 4 ครั้งใน 1 ภาคการศึกษา ผู้วิจัยตรวจสอบความตรงโมเดลสมมติฐานซึ่งเป็นข้อมูลระยะยาวโดยใช้การวิเคราะห์พหุระดับ วิเคราะห์ข้อมูลด้วยโปรแกรม Mplus 8.0 ผลการวิจัยระบุว่า ร้อยละ 32.6% ของความแปรปรวนมาจากความแตกต่างระหว่างบุคคล (Intraclass coefficient = .326) คะแนนเริ่มต้นของการให้เหตุผลเชิงคุณธรรมมีความแปรปรวน ($\sigma^2_{\mu 0}$ = 2.719, $p < .001$) ความแปรปรวนของคะแนนการให้เหตุผลเชิงคุณธรรมภายในตนเองสามารถอธิบายได้ด้วยเวลาและความอ่อนไหวเชิงคุณธรรม คะแนนการให้เหตุผลเชิงคุณธรรมจะเพิ่มขึ้น 0.19 หน่วย ในการวัดครั้งต่อไป ($p < .05$) เมื่อพิจารณาความแปรปรวนระหว่างบุคคล ค่าเฉลี่ยคะแนนการให้เหตุผลเชิงคุณธรรมเริ่มต้นมีค่าเท่ากับ 11.33 หน่วย ($p < .01$) และที่คะแนนเริ่มต้นพบว่า นักเรียนหญิงมีคะแนนสูงกว่า นักเรียนชาย 1.377 หน่วย ความรู้ในเนื้อหาเทคโนโลยีดีเอ็นเอและผลสัมฤทธิ์ทางการเรียนวิชาชีววิทยาก่อนหน้าพบว่า ไม่สามารถอธิบายความแปรปรวนภายในและระหว่างบุคคลได้ตามลำดับ งานวิจัยนี้ได้ให้ข้อเสนอแนะการจัดการศึกษาเพื่อส่งเสริมคุณธรรม

คำสำคัญ : การประยุกต์ใช้เทคโนโลยีชีวภาพสมัยใหม่, โมเดลโค้งพัฒนาการ, การให้เหตุผลเชิงจริยธรรม

[1] ภาควิชาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพฯ 10900
onrumpa.ku@ku.th

[2] *ผู้รับผิดชอบบทความหลัก,*
ภาควิชาการศึกษา คณะศึกษาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพฯ 10900
pongprapan.p@ku.th

[3] ภาควิชาพันธุศาสตร์ คณะวิทยาศาสตร์ มหาวิทยาลัยเกษตรศาสตร์ กรุงเทพฯ 10900
fsciuwa@ku.ac.th

Introduction

The application of modern biotechnology raises considerable moral questions. Therefore, from an educational perspective, many science educators (Pedretti, 1999; Zeidler, 1984) have recommended that when dealing with such topics—especially gene and DNA technology—morality and ethics must be taken into account. If we seriously consider the proposition that the science curriculum must prepare future citizens (AAAS, 1990), then the moral dilemmas posed by the applications of modern biotechnology can provide opportunities for science classrooms to prepare students to be functional citizens who can appropriately apply their understanding of science in making well-informed decisions (Dawson & Schibeci, 2003).

Many science educators have advocated enhancing moral reasoning in science classrooms through the use of socioscientific issues (SSIs), especially genetics-related dilemmas (Andrew & Robotton, 2001; Evans, 2002; Stock & Campbell, 2002; Zeidler, 1984; Zohar & Nemet, 2002). SSIs features complex real-world problems that remain unresolved. SSI-based teaching can stimulate the consideration of moral issues and implications for students across various age groups, including middle school, high school, and university (Grace & Ratcliffe, 2002).

Time is required to develop the ability of moral reasoning (National Research Council of Thailand, 2013; Piaget, 1932). Therefore, the change in students' moral reasoning over time needs to be systematically captured using a longitudinal data analysis framework. To capture developmental changes over time in a targeted construct (in this case, moral reasoning), many researchers have advocated the use of multilevel analysis or hierarchical linear modeling (HLM). Longitudinal data, or repeated measures data, can be viewed as multilevel data with repeated measurements nested within individuals (Singer & Willett, 2003). HLM is used for modeling within-person systematic changes (i.e., intrapersonal level) and between-person differences (i.e., interpersonal level) in developmental outcomes across different measurements over time (McCoach & Kaniskan, 2010; Singer & Willett, 2003). In addition, HLM does not require the repeated measurement to be equally spaced, and it can effectively handle missing data by treating each testing occasion of an individual participant as a separate case. Further, HLM can handle both continuous and categorical predictors (Hox, 2010; Raudenbush & Bryk, 2002).

In this light, this study aims to answer the following research questions: (1) How do the overall moral reasoning scores change over the course of the intervention? (2) How large are students' within- and between-person variations of the moral reasoning scores? and (3) What are the factors that can explain students' between- and within-person variations of the moral reasoning scores? Based on our synthesis of the relevant literature, we hypothesize the following variables as potential determinants to explain such variations: (1) moral sensitivity, (2) conceptual understanding of basic genetics and DNA technology concepts, (3) previous achievement in basic biology, and (4) student's gender.

In a longitudinal data analysis framework, some of the abovementioned variables are time-varying, that is, they change and affect the variation of the measurement occurrence of moral reasoning within a subject over time. We hypothesized that these variables are moral sensitivity and conceptual understanding. By contrast, other variables are time-invariant. We thought that some variables were already innate in a subject (e.g., gender, previous biology achievement (before intervention). This type of predictors would explain either the variance of the initial status and rate of change between the subjects.

Development of Moral Reasoning in Science Teaching

As we wish to enhance students' moral reasoning, we synthesized the essential features of teaching bioethics and incorporated them into a well-established SSI approach (Zeidler, 1984). Underlying these strategies are the assumptions that learning must be made meaningful to students and should be situated and dialogic. We reviewed many interventions in school science contexts targeting bioethics and moral reasoning. We found that most relied on Kohlberg's framework (1969), Rest's framework (1986), and the extensive works of Sadler and Zeidler (2005a). The five essential features of teaching biotechnology that we used to guide the design and construction of our instructional modules are listed below:

1. Selection of a provoking issue that connects the student's experience and science content.

2. Have students encounter different views and engage in a critical discourse.

3. Encourage students to express and justify their standpoint in a respectful and supportive environment.

4. Train students to monitor and regulate their moral learning.

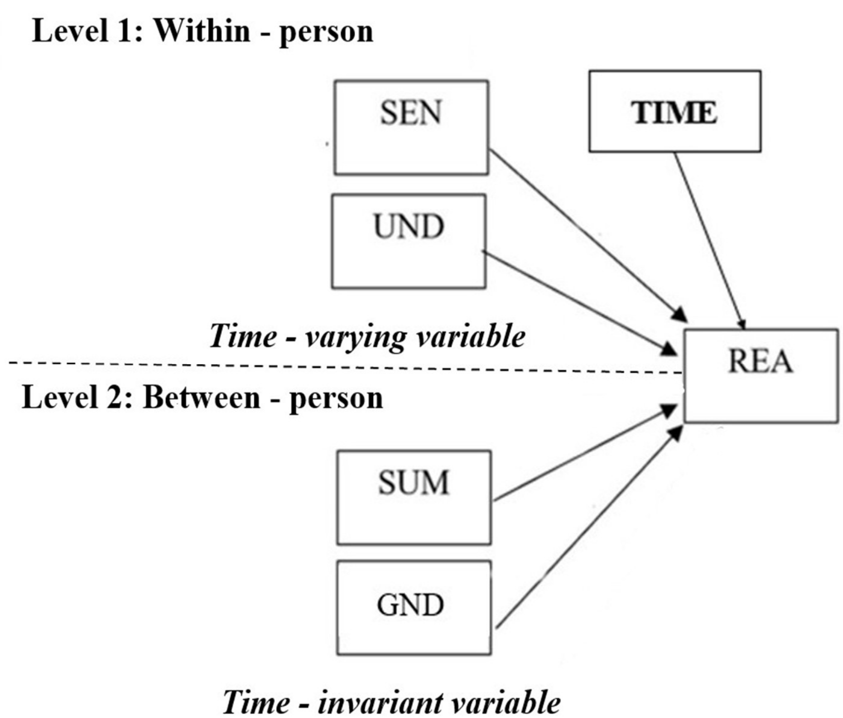
5. Teachers as facilitator.

Hypothetical Growth Model

In growth analyses, the initial status, rate of increase, and shape of the growth trajectory represent the primary parameters of interest. In addition, many growth models strive to explain the factors that influence individual variations in development (Hox, 2010). Our hypothetical model aims to illustrate factors that may help to explain the development of high school students' moral reasoning. Based on the growth curve modeling within a multilevel framework (Hox, 2010; Singer & Willett, 2003), this investigation constructed a two-level hypothetical model to explain students within- and between-person variations for the outcome variable, moral reasoning (REA). Moral reasoning refers to the logical process of determining whether an action is right or wrong. At level 1, the predictors include time-varying variables: moral sensitivity (SEN), understanding of DNA-Technologies (UND), and occasion of measurement (TIME). Moral sensitivity is the ability to identify the existing moral problem and understand the moral consequences of the decisions made on the patient's part. The understanding of DNA-Technologies is students' conceptual understanding of key concepts in DNA Technology Unit in Biology subject. At level 2, the predictors were time-invariant variables that are constant over the time, including previous biology achievement (SUM) which is academic outcomes that indicate the extent to which a student has achieved their learning Biology and student gender (GND) (Figure 1). This two-level conceptualization implies the need for a model in which the parameters in the first level become the outcome variables in the second level. Because of this multilevel character, it is convenient to refer to this conceptualization of growth curve modeling as a hierarchical linear modeling or HLM.

Figure 1

A Hypothetical Growth Curve Model of Moral Reasoning on the Application of Modern Biotechnology.



Upon applying the above model to our study, at level 1, we obtain

$$REA_{ti} = \pi_{0i} + \pi_{1i}\, TIME_{ti} + \pi_{2i}\, SEN_{ti} + \pi_{3i}\, UND_{ti} + \varepsilon_{ti} \quad (1)$$

The coefficients at level 1 are often indicated by the Greek letter π. The subject-coefficients at the second level can be represented by the usual Greek letter γ. In Eq. (1), REA_{ti} is the moral reasoning score of individual *i* measured at measurement occasion *t*, $TIME_{ti}$ is the time variable that indicates the measurement occasion, and SEN_{ti} and UND_{ti} are time-varying covariates. The second-level predictors including SUM and GND are time-invariant covariates. They are hypothesized to explain the variation of the initial status (π_{0i}) and the rate of growth (π_{1i}), respectively.

$$\pi_{0i} = \gamma_{00} + \gamma_{01}\, SUM_i + \gamma_{02}\, GND_i + \mu_{0i}$$

$$\pi_{1i} = \gamma_{10} + \gamma_{11}\, SUM_i + \gamma_{12}\, GND_i + \mu_{1i} \quad (2)$$

where ε_{ti} represents the residual of person *i*'s score at time *t* from his/her model predicted score at that time point. The person-level residuals μ_{0i} and μ_{1i} represent the deviation of person *i*'s intercept and slope from the overall intercept and slope.

Method

Population and Sample

The population in this study are grade 12 students studying DNA technology unit in academic year of 2017 in Saraburi province. There were, in total, 21 high schools in Saraburi. Three schools with different sizes (extra-large, large, and medium) were randomly selected. The sample were 206 Grade 12 students. All of them had learned fundamental and prerequisite genetics concepts before our intervention. These students were taught by three biology teachers who volunteered to participate in this study. Each taught two intact biology classes. The teaching method used was SSI-based instruction emphasizing case-based discussion. The cases were issues and concerns on the application of DNA technology. After introducing the case, the students engaged in either a whole-class discussion or, where appropriate, were divided into pro and con groups based on their opinion and discussion.

Interventional bioethics modules

The teaching intervention comprised five modules in the units of Molecular Genetics and DNA Technology in high school biology. We named our intervention, Bioethics for Informed-Decision Modules or BID modules. The teaching modules integrated SSI-based teaching. We wished to enable students to develop moral reasoning skills so they could move beyond a “gut reaction” to more informed and critical positions. The goals of BID modules were to encourage them to ask moral questions, understand moral principles, and apply them and scientific understanding to make informed decisions considering various stakeholders' perspectives on biotechnological issues. We tried to cultivate critical reasoning skills, moral values, and leadership in "hard" topics in Biology. BID comprised five modules namely, XYY syndrome born to be bad, Gene therapy and Huntington's disease, GM mosquito, Alzheimer's disease and the secret letter, and Thyroid cancer and genetic diagnosis. The implementation of the BID modules can be flexible depending on school contexts. We only designed classroom activities, worksheets, and materials and gave teachers instructional guidelines. In each module, we provided scenarios about moral dilemmas in the application of DNA technology. These were fictional or realistic and were conducted after teaching science content. The students engaged in structured discussion in a safe learning environment.

Instruments and Measures

We used two research instruments. The first instrument was a questionnaire measuring moral sensitivity and moral reasoning. It was adapted from the work of Clarkeburn (2002) and Fowler et al. (2009), and we created some more scenarios related to the Thai context; "Golden Rice" and "BT cotton." The questionnaire comprised four open-ended questions covering the application of DNA technology on health, economy, society, and environment. The instrument began with a short paragraph introducing a dilemma that was complex and multifaceted and that had to do with balance and trade-offs. To measure sensitivity, the students were asked to list questions or concerns they had after reading the text. Regarding moral reasoning, with the same dilemma, students were asked further to give their position and justification for their choice. Their responses were interpreted and classified into one of predetermined ordered categories. The second instrument measures the conceptual understanding of molecular genetics and DNA technology. It was a multiple choices test comprising 12 items. The item difficulty ranged between .3 and .7. The item discrimination, the items fall in acceptable values are .2 or higher.

The moral sensitivity, a time-varying variable, had a total score of 12 points using a scoring rubric adapted from Fowler et al. (2009). The outcome variable, moral reasoning (REA), had a total score of 20 points. We adapted a five-level scoring rubric based on Črne-Hladnik et al. (2012) to assess moral reasoning. The understanding of DNA-technology-related concepts was measured from a multiple-choice test with a maximum score of 12 points.

All instruments were content validated by experts and pilot tested. After the pilot test, interviews were conducted with some students to gauge whether and how they understood the dilemma and questions and which items caused them the most trouble. Based on results from the pilot test, the instruments were revised. The data of other variables in the model, such as gender and previous biology achievement, were retrieved from the students' academic record provided by their school.

Data Collection

This study was conducted in the first semester of the 2017 academic year. Each student was assessed for the outcome variables and time-varying predictors four times over the course of the intervention. The first data collection occasion occurred at the beginning of the semester before the implementation of the intervention (TIME 0); the second occasion, after the first module (TIME 1); the third occasion, after the third module (TIME 2); and the fourth occasion, at the end of the semester (TIME 3).

Data Analysis

We described the data of all variables in the model using descriptive statistics and then conducted HLM to examine and explain students' within- and between-person variation (Raudenbush & Bryk, 2002). We estimated a series of multilevel models using Mplus version 8.0 software. To conduct a multilevel analysis, we adopted the multi-step approach suggested by Hox (2010) and Singer and Willett (2003). By using the multilevel framework, we could explore intraindividual changes over time (level 1) and individual differences in the nature of the change (level 2). To explain the variation, the predictors were added to each level at the time. The modeling was run in sequence from the simplest, baseline model, to the most complex, full model. A previous model determines and guides the specification of a subsequent model. The variations of the outcome variables were expected to be reduced and explained in the later models. In this study, there are five sub-models, as described in order below.

Model A, the intercept-only model, is a model with no predictors at either level. This model contains only an intercept and error terms at the occasion and the subject levels. Model A will help partition the total outcome variation. From this model, we calculate the intraclass correlation (ICC, ρ) that quantifies the proportion of the observed variation in the outcome that was attributable to the effect of clustering within and between students. It was calculated as the ratio of the group-level error variance to the total error variance where $\sigma_{\mu 0}^2$ is the variance of the level-2 residuals and σ_e^2 is the variance of the level-1 residuals as follows:

$$\rho = \frac{\sigma_{\mu 0}^2}{\sigma_{\mu 0}^2 + \sigma_e^2}$$

In Model B, the unconditional growth model, the time variable was added as a linear predictor with the same coefficient for all subjects at the intrapersonal level. We used a measurement occasion variable (TIME) to explain the occasion-level variance in the outcome variable (REA) over the course of the intervention. To examine how much the variance decreased when explanatory variables were added in the subsequent models, we used Model B as a baseline model as suggested by Hox (2010). The intercept represents the value of REA_{ti} when time = 0 or a person i's initial status. At level 2, this model shows between-person variability in terms of the initial status and the linear growth rates.

Model C had a random intercept and a slope with a level-1 predictor model. We added the time-varying covariates UND and SEN to this model at level 1 to explain the within-person variance.

Model D, the inter-individual level model, was used to explain the between-person variability of the initial status. We introduced time-invariant predictors, GND and SUM, to explain the variability of the initial stage of moral reasoning between students at level 2.

Model E, the inter-individual level model, was used to explain the between-person variability of the rate of change. We introduced time-invariant predictors, GND and SUM, to explain the variability of the rate of change in the development of moral reasoning between students. Notably, Models D and E would proceed only if there was a significant amount of variability of initial status and rate of change between students.

To determine the effect of the significant predictors, we calculate the proportion of variance reduced (explained) as

$$\text{the proportion of variance reduced} = \frac{(\text{the error in a baseline model}) - (\text{the error in a subsequent model})}{\text{the error in a baseline model}}$$

Results

Initially, we had 206 participants. However, of these, 16 showed no variation in their responses over time and were therefore removed from the dataset. Therefore, the effective sample size was 190 students. We examined the distribution of the outcome variable moral reasoning (REA) in the disaggregated data file with 190 × 4 = 760 observations with 47 missing data. This section first presents the descriptive analysis results (Table 1) and then presents the multilevel growth curve modeling results (Table 2).

Table 1 presents the descriptive statistics of the outcome variable and level-1 and level-2 predictors. The average score of the moral reasoning outcome variable was 14.17. The average score of the level-1 explanatory variables, SEN and UND, was 7.23 and 6.95, respectively. The average score of the level-2 explanatory variable, SUM, was 136.25. Skewness assesses the extent to which a variable's distribution is symmetrical. As a general guideline, a skewness value between −1 and +1 is considered excellent. Kurtosis is a measure of whether the distribution is too peaked. the general guideline, when both skewness and kurtosis are close to zero, the pattern of responses is considered a normal distribution" (Hair et al., 2022, p. 66). In this study, the skewness and kurtosis of the variables were less than 1 and close to 0 respectively, suggesting that all variables were symmetrical and normally distributed.

Table 1

Descriptive statistics of each variable

Score			N (%)	Max	Min	Mean	SD	Skewness	Kurtosis
Outcome	REA			20	8	14.17	2.879	-.305	-.708
Level-1 explanatory variables	SEN			12	4	7.23	1.849	.232	-.426
	UND			12	0	6.95	3.057	-.168	-1.082
	TIME			-	-	-	-	-	-
Level-2 explanatory variable	SUM			177	99	136.25	13.654	.401	.088
	GND	Total	190 (100)	-	-	-	-	-	-
		Female = 1	122 (64.2)						
		Male = 0	68 (35.8)						

Note: REA (moral reasoning), SEN (moral sensitivity), UND (understanding of DNA technologies), TIME (occasion of measurement), SUM (previous biology achievement), GND (student gender)

Next, we present the results obtained from the HLM analysis based on the approach by Hox (2010) and Singer and Willett (2003). Table 2 lists the results of the parameter estimates for the growth models. Each model is described in detail below.

The result of Model A indicates that the intercept (Υ_{00}), which is the average REA across all individuals and occasions without predictors, is 14.195. The variance of the repeated measurement-level residual errors, denoted by σ_e^2, is estimated as 5.572. The variance of the student-level residual errors, denoted by $\sigma_{\mu 0}^2$, is estimated as 2.704. The ICC or the proportion variance at the second level, calculated as ρ = $\sigma_{\mu 0}^2$ / ($\sigma_{\mu 0}^2$ + σ_e^2), is .326. In other words, 32.6% of the total variance is contributed by individual differences, and the remaining variance is due to intrapersonal variations. The fit indices AIC and BIC provide the model fit information. When explanatory variables are added to the model, the AIC and BIC are expected to decrease.

In Model B, the time variable was added as a linear predictor. The variance of the repeated measurement-level residual errors, denoted by σ_e^2, is estimated as 5.505. The variance of the student-level residual errors, denoted by $\sigma_{\mu 0}^2$, is estimated as 2.719. The ICC or the proportion variance at the second level, calculated as ρ = $\sigma_{\mu 0}^2$ /

($\sigma^2_{\mu 0}$ + σ^2_e), is .330. This shown that, 33% of the total variance is contributed by individual differences. In other words, about 70% of the remaining variance is due to intrapersonal variations. The result of Model B predicts a value of 13.862 (Υ_{00}) at the first occasion, and it increases by .198 (Υ_{10}) on subsequent measurement occasions. According to this model, students have a positive rate of change in their moral reasoning score. Model B also shows a statistically significant variation in the intercept within students in the population (σ^2_e = 5.505; p < .001).

The result obtained from a revised version of Model C after trimming insignificant predictors indicated that only moral sensitivity (SEN) was a significant predictor. The conceptual understanding of DNA technology (UND) cannot explain the variance of occasion level. The value of the moral reasoning means score (Υ_{00}) is 12.221 (p < .001) at the first occasion, and it increases by 0.189 (Υ_{10}) on subsequent measurement occasions. The regression coefficient for SEN is .229 (Υ_{11}); this means that for each unit increase in scale point on the ethical sensitivity measurement, the ethical reasoning is expected to increase by .229 scale points in succeeding measurements after controlling the other significant predictor (TIME). The fit indices AIC and BIC decrease.

In Model D, all level-2 predictors were added to explain the variation of the initial status of moral reasoning between students. After trimming the insignificant predictor; biology achievement (SUM), gender (GND) had a regression coefficient of 1.376 (p < .001), implying that a gender difference existed in moral reasoning at the initial stage. The gender can reduce and explain the variation of the initial status between students by ((2.719–1.933)/2.719) = .29 or 29%. At the initial point, girls scored 1.377 points more than boys on the moral reasoning measure.

Model E was the same as Model D and was added to explain the variability of the rate of change between students. The result of Model E indicated that the variance of the slope (TIME) between a person was insignificant (0.163, p = .069), implying that all students have the same rate of change in moral reasoning score over the course of the instruction. Thus, we revised Model E by fixing the variance of slope to 0. Finally, we reran the revised Model E and found that all students have the same positive rate of change of moral reasoning score at 0.19 unit. An evaluation of the fit indices AIC and BIC indicated that the revised Model E was the best model to explain the within-student variance of the measurement occasions and the between-student variance of the initial status of moral reasoning.

Table 2

Parameter Estimates of Submodels in the Multilevel Growth Curve Modeling

	Parameter Estimates					
	Model A (Null model)	**Model B** (+TIME)	**Revised Model C** (+only SEN)	**Revised Model D** (+only SEX)	**Revised Model E**	**Standardized Coefficients**
Level 1: Within level						
REA on TIME (Υ_{10})		0.198**	0.189*	0.190*	0.190*	**.15**
REA on SEN (Υ_{11})			0.229**	0.228**	0.228**	**.07**
REA on UND (Υ_{12})			Trimmed			
Residual variances REA (σ^2_e)	5.572**	5.505**	5.544**	5.529**	5.529**	
Level 2: Between level						
Means REA (Υ_{00})	14.159**	13.862**	12.221**	11.330**	11.330**	
REA on SUM (Υ_{13})				Trimmed		
REA on GND (Υ_{14})				1.376**	1.377**	**.69**
Residual variance ($\sigma^2_{\mu0}$)	2.704**	2.719**	2.334**	1.933**	1.934**	
Means of Slope					0.190*	
Variance of Slope					Fixed at 0	
Model fit information						
Number of free parameters	3	4	5	6	6	
Loglikelihood	-1722	-1718.8	-1717.9	-1700.6	1700.57	
AIC	3450.06	3445.66	3433.8	3413.13	3413.13	
BIC	3463.76	3463.23	3456.65	3440.55	3440.55	
ICC	.326	.329	.288	.292		

* $p < .05$, ** $p < .01$

The coefficients in Table 2 were all unstandardized regression coefficients. To interpret them properly, standardized regression coefficients were calculated from revised Model E using the equation below (Hox, 2010). This will help identify the strongest or the weakest explanatory variable that explains the outcome variable.

The result of the standardized regression coefficients indicated that gender (GND) has the largest impact (β = .69), followed by time (β = .15) and moral sensitivity (β = .07). The overall results from unstandardized parameters indicate that the initial score of moral reasoning for male students is 11.33 points. The initial score of female students is 1.377 points higher than that of male students. Their score increases constantly by 0.19 points in each subsequent measurement. Taking into control of the previous variables, with each unit increase in the scale point for moral sensitivity, the ethical reasoning is expected to increase by 0.228 points.

Discussion

Based on the ICC from Model B, we found that around 70% of the variance arises from intrapersonal differences in the repeated measure. This may be influenced by personal experience of an issue, family biases, or cultural values (Sadler & Zeilder, 2005b). Individuals are posited to respond to moral situations based on their own unique perceptions, which make their action choices dynamic and unpredictable.

In Model B, when a linear time variable was added as a predictor to the measurement occasion level, we found that it was significant and positive. This indicates that over the course of instruction, moral reasoning increases in every subsequent measurement. This means that the moral qualities of our subject can be fostered or hampered by given learning experiences and opportunities in their classroom environment. Hart (2005) asserted that the development of a moral identity was nurtured by prolonged learning experiences related to self-awareness, continuity through time and place, the self in relation to others, and the self as a basis for strong evaluations. In our case, we capture the growth of students' moral reasoning over four months. They included the important plans, goals, and values that form a basis for individual's perceiving, judging, and acting. These kinds of experiences are consistent with the essential features of teaching bioethics underlying BID modules.

We, therefore, discuss the findings from only Model E. We found a weak positive relationship between the moral sensitivity and the moral reasoning score. The minimal effect may be explained by the issues used in the test. Some examples like a GM cow for patients with cystic fibrosis are new and are unrelated to personal experiences. Further, genetic disorders are not widely known in Thai society. Therefore, students might not have been impacted by this issue. The students may have felt less sensitive when exposed to the situation.

We found that the conceptual understanding of the issue and previous biology achievement are not significant predictors. This finding is consistent with some previous studies. Dawson and Taylor (1999) and Olsher and Dreyful (1999) noted that ethical decisions did not differ between nonscience and science major students. Additionally, previous studies have documented differences in argumentation quality. Sadler and Zeidler (2005b) indicated that subjects who knew a great deal about science and genetics and those who knew little displayed similar patterns in terms of their reliance on rational, emotive, and intuitive arguments with respect to genetic engineering issues.

To explain the variance in the initial status of moral reasoning between the students, we found that gender is the only significant level-2 predictor, and it has the largest impact among all predictors in the final model. Female students have a higher score on moral reasoning than their male counterparts. Gilligan (1993), a psychologist, explained that women tend to consider moral issues in terms of relationships, caring, and compassion, whereas men are likely to consider such issues in terms of justice, rules, and individual rights. Dawson (1992) similarly noted that in solving ethical dilemmas, females are likely to primarily respect feelings whereas males are likely to primarily respect rights.

The growth in moral reasoning over the course of BID modules may have caused by the essential features of teaching bioethics, including an introduction to the dilemmas and an open discussion based on reliable resources. It consists of structured discussions and ethical inquiry activities based on situated and dialogic learning. These activities emphasize ethical deliberation to enable students to face the implications of modern biotechnology from multiple perspectives, including human health, environment or ecosystem, and economics. Consistent with Sadler and Zeidler (2005a), moral reasoning that ignores real-world evidence is fundamentally flawed. Science classrooms that deny emotive venues of discourse in the discussion of social-science issues curtail student's moral development.

In the final model, after entering all significant predictors to both levels, we could not reduce the occasion measurement variance (level-1) in the baseline model with the trivial effect of moral sensitivity. By contrast, we could reduce and explain around 29% of the variance between students (level-2) with gender. This suggests that the rest of the variance within and between students is unexplained. We must have omitted some other significant predictors from our hypothetical model. Some potential variables that may help explain the variance include, but are not limited to, the subjective norm, normative belief, perceived behavior control, religiosity, and obedience to the law. These will be investigated in further research.

Limitation and Suggestion

We believed that the advantages offered by this longitudinal study research design could help science teachers capture students' moral growth. We confirmed that classroom assessment should be ongoing and reflective and should inform teachers' teaching practice. This guides future studies to extend the course of the intervention with a wider space between waves of observations. This would help to capture a holistic picture of students' moral reasoning growth. Further, we might use other advanced statistical analyses, for example, latent growth curve modelling, to explore other factors that might have an impact on students' moral reasoning construct.

Implication

Even though growth in moral reasoning may not occur obviously in one semester, science teachers should not ignore it in their science classroom. With time and continuous practice, students will show better results. Science educators need to move students beyond their initial reactions, not as a means of necessarily changing these views, but as a means for encouraging critical reflection. Students are expected to not only make reasoned judgments of scientific data; the real-world challenges them to consider what is right, and this necessarily entails normative ethical thoughts like "ought" and "should." Thus, policymakers or stakeholders should better prepare teachers to teach the moral dimension of science in science classrooms.

References

American Association for the Advancement of Science (AAAS). (1990). *Science for All Americans*. Oxford University Press.

Andrew, J., & Robottom, I. (2001). Science and ethics: Some issues for education. *Science Education, 85*, 769–780. https://doi.org/10.1002/sce.1038

Bryk, D.T., & Raudenbush, S. W. (1992). *Hierarchical linear model: Applications and data analysis methods.* Sage.

Clarkeburn, H. (2002). A test for ethical sensitivity in science. J*ournal of Moral Education, 31*(4), 439–453. https://doi.org/10.1080/0305724022000029662

Črne-Hladnik, H., A. Hladnik, B. Javornik, K. Košmelj and C. Peklaj. (2012). Is judgment of biotechnological ethical aspects related to high school students' knowledge? *International Journal of Science Education, 34*(8), 1277–1296. https://doi.org/10.1080/09500693.2011.572264

Dawson, L. M. (1992). Will feminization change the ethics of the sales profession? *Journal of Personal Selling and Sales Management*, 21–32. https://www.jstor.org/stable/40471451

Dawson, V. M. & Taylor, P.C. (1999). Teaching bioethics in science: Does it make a difference? *Australian Science Teachers' Journal, 45*, 59–64.

Dawson, V. M., & Schibeci R. A. (2003). West Australian school students' understanding of Biotechnology. *International Journal of Science Education, 25*, 57–69. https://doi.org/10.1080/09500690210126720

Evans, J. H. (2002). *Playing God?: Human genetic engineering and the rationalization of public bioethical debate.* University of Chicago Press.

Fowler, S. R., Zeidler, D. L., & Sadler, T. D. (2009). Moral sensitivity in the context of socioscientific issues in high school science students. *International Journal of Science Education, 31*(2), 279–296. https://doi.org/10.1080/09500690701787909

Gilligan, C. (1993). *In a different voice: Psychological theory and women's development.* Harvard University Press.

Grace, M. M., & Ratcliffe, M. (2002). The science and values that young people draw upon to make decisions about biological conservation issues. *International Journal of Science Education, 24*, 1157–1169. https://doi.org/10.1080/09500690210134848

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (3rd ed.).* Sage.

Hart, D. (2005). The Development of Moral Identity. In G. Carlo & C. P. Edwards (Eds.), *Moral motivation through the life span* (pp. 165–196). University of Nebraska Press.

Hox, J. (2010). *Multilevel analysis: Techniques and applications.* Routledge.

Kohlberg, L. (1969). Stage and sequence: The cognitive developmental approach to socialization. In D. Goslin (Ed.), *Handbook of socialization theory and research* (pp. 347–480). Rand McNally.

McCoach, D. B., & Kaniskan B. (2010). Using time-varying covariates in multilevel growth models. *Frontiers in Psychology, 1*, 1–12. https://doi.org/10.3389/fpsyg.2010.00017

National Research Council of Thailand. (2013). *Teachers and strengthening ethics to students*. Bangkok. [in Thai]

Olsher, G., & Dreyful, A. (1999). The 'ostension-teaching approach as a means to develop junior-high student attitudes towards biotechnologies. *Journal of Biological Education, 34*, 24–30. https://doi.org/10.1080/00219266.1999.9655679

Pedretti, E. (1999). Decision making and STS education: Exploring science knowledge and social responsibility in schools and science centers through an issues-based approach. *School Science and Mathematics, 99*, 174–181. https://doi.org/10.1111/j.1949-8594.1999.tb17471.x

Piaget, J. (1932). *The moral judgment of the child.* Harcourt, Brace.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Application and data analysis methods (2nd ed.).* Sage.

Rest, J. R. (1986). *Moral development: Advances in research and theory.* Praeger.

Sadler, T. D., & Zeidler, D. L. (2005a). Patterns of informal reasoning in the context of socio-scientific decision-making. *Journal of Research in Science Teaching, 42*(1), 112–138. https://doi.org/10.1002/tea.20042

Sadler, T. D., & Zeidler, D. L. (2005b). The significant of content knowledge for informal reasoning regarding socioscientific issues: Applying genetics knowledge to genetic engineering issues. *Science Education, 89*(1), 71–93. https://doi.org/10.1002/sce.20023

Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis.* Oxford University Press.

Stock, G., & Campbell, J. (2000). *Engineering the human germline: An exploration of the science and ethics of altering the genes we pass to our children.* Oxford University Press.

Zeidler. D. L. (1984). Moral issues and social policy in science education: Closing the literacy gap. *Science Education, 68,* 411–419. https://doi.org/10.1002/sce.3730680406

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching, 39*, 35–6. https://doi.org/10.1002/tea.10008