From the
AERA Online Paper Repository
http://www.aera.net/repository

Paper Title  How to Make Moral Education More Effective: From a Brain Study to Policy Making
Author(s)  Hyemin Han, The University of Alabama; Stephen J. Thoma, The University of Alabama; Firat Soylu, The University of Alabama; Kangwook Lee, KAIST
Session Title  Moral Development and Education: Knowledge to Action
Session Type  Paper
Presentation Date  4/27/2017
Presentation Location  San Antonio, Texas
Descriptors  Adolescence
Methodology  Qualitative
Unit  SIG-Moral Development and Education
DOI  10.302/1172082

Each presenter retains copyright on the full-text paper. Repository users should follow legal and ethical practices in their use of repository material; permission to reuse material must be sought from the presenter, who owns copyright. Users should be aware of the AERA Code of Ethics.

Citation of a paper in the repository should take the following form:
How to Make Moral Education More Effective? From a Brain Study to Policy Making

Hyemin Han  Stephen J. Thoma  Firat Soylu

University of Alabama

Kangwook Lee

KAIST

Abstract

This paper suggests how to develop more effective moral educational programs by utilizing interdisciplinary research methods including neuroimaging, social psychological intervention, evolutionary modeling and deep learning methods. Our interdisciplinary research program consists of three steps: identification of core psychological processes involved in moral functioning at the neural level, development of small-scale intervention programs tweaking the identified psychological processes, prediction of long-term, large-scale outcomes of designed interventions using computational methods. We discuss how this research program will inform moral educators and educational policy makers who wish to develop effective moral educational programs in larger educational settings in the long term with actual neuroimaging and intervention experimental data.

Introduction

Recently, neuroscientists have conducted various experiments focusing on the neural correlates of human morality (Bzdok et al., 2012; Sevinc & Spreng, 2014). The findings have enhanced our understanding of the biological mechanism and substructure of moral functioning with evidence. However, it is still not clear how to apply the neuroscientific findings to moral education in practice and how to scale up the moral educational programs developed by psychological experiments in a lab in order to make them effectively work in a large scale, such
as a school or district level. Thus, we consider how to establish an interdisciplinary research program embracing neuroscience, experimental psychology, educational practice and policy making that can contribute to the improvement of moral educational in the reality (Han, 2016a). First, we discuss how neuroimaging methods including meta-analysis and functional magnetic resonance imaging (fMRI) can identify central neural mechanisms associated with moral functioning of interest, which will be targeted by moral educational interventions. Second, we demonstrate psychological experiments inspired by social psychology that are designed to target the psychological mechanisms identified by neuroimaging. Third, we discuss how to apply evolutionary modeling and computer simulation methods in order to develop appropriate interventions and inform educational policy makers. We demonstrate how the research program with these three procedures can produce applicable outcomes, the application of stories of moral exemplars in educational settings intending the promotion of moral motivation.

**Utilizing Neuroimaging Methods**

Neuroimaging methods can potentially contribute to the expansion of our knowledge regarding how human morality is functioning with biological evidence, which could not be approached through traditional research methods, such as paper-and-pencil survey (Ito & Cacioppo, 2007; Kristjánsson, 2013). However, how these scientific findings can eventually contribute to the development of moral education in practice is still unclear (Han, 2016b; Kristjánsson, 2015). Therefore, it would be necessary to consider how to connect the factual findings from neuroimaging studies to educational application in practice by establishing a valid interdisciplinary research program. As a first step in this process, we consider two specific neuroimaging methods, i.e., meta-analysis and fMRI, which can illuminate psychological processes involved in moral functioning, as components in the research program.
A meta-analysis of previously conducted neuroimaging studies can identify common activation foci of psychological functions of interest, moral functioning in case of our study. It can address several issues associated with traditional neuroimaging methods, such as the lack of statistical power originating from relatively small sample sizes, idiosyncrasies in experimental designs (Costafreda, 2009; Etkin & Wager, 2007; Wager, Jonides, & Reading, 2004) and possibility of reverse inference in interpretation (Poldrack, 2011). Several meta-analyses of previous neuroimaging studies focusing on moral functioning have demonstrated common activation foci associated with moral psychological processes (Bzdok et al., 2012; Han, 2017; Sevinc & Spreng, 2014). The common findings obtained by these meta-analyses provide a foundation for hypothesis setting in future fMRI studies (Han, 2016c). Taken together, the finding of the meta-analyses across diverse morality-related task conditions (see Fig. 1) found activation in brain regions associated with self-related processes, particularly autobiographical self and self-evaluation, that is, the default mode network (DMN) and cortical midline structures (CMS). Given these results, it is plausible that selfhood seems to commonly engage with moral functioning. Therefore, it would seem profitable to focus the next generation of fMRI experiments on examining psychological processes that will be targeted during intervention experiments.
Figure 1. Common activation foci of moral functioning extracted from the meta-analysis.

In addition to the meta-analyses, fMRI experiments have demonstrated the neural correlates of moral function within diverse experimental paradigms (Greene, Nystrom, Engell, Darley, & Cohen, 2004; Robertson et al., 2007). Particularly informative is, a recent fMRI experiment testing hypotheses based on the findings from the previous meta-analyses (Han, Chen, Jeong, & Glover, 2016). This fMRI experiment investigated how brain regions associated with selfhood, the DMN and CMN, moderated activity in other brain regions associated with moral emotion and motivation, such as the insula, while solving moral problems as assessed using the psychophysiological interaction analysis and Granger causality analysis methods (Friston et al., 1997; Seth, Barrett, & Barnett, 2015). Fig. 2 demonstrates the results of the analysis. As hypothesized, the analysis indicated that neural activity in regions associated with selfhood in the DMN and CMS, particularly the medial prefrontal cortex and posterior cingulate cortex, significantly moderated activity in moral emotion and moral motivation-related regions,
particularly the insula. Consequently, the fMRI experiment was able to support the hypotheses based upon previously published neuroimaging studies of morality and their meta-analyses and to identify psychological processes that will be targeted by intervention experiments, that is self-related psychological processes.

Figure 2. fMRI experiment results. (a) Psychophysiological analyses. (b) Granger causality analysis.

**Psychological Intervention Methods**

Interventions based on psychology, particularly social and educational psychology, have improved students’ academic achievement and social adjustment in diverse educational settings (Cohen & Garcia, 2008; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009; Yeager, Trzesniewski, & Dweck, 2013; Yeager, Trzesniewski, Tirri, Nokelainen, & Dweck, 2011; Yeager & Walton, 2011). Thus, developed psychological intervention methods can provide
useful insights about how to design more effective moral education programs. Basically, psychological interventions are designed to tweak psychological processes that are fundamentally associated with the targeted developmental outcome (Miller & Prentice, 2012). Based on the purpose and nature of intervention research, two psychological intervention experiments using the stories of moral exemplars were conducted to test which type of exemplary stories better promoted motivation to engage in moral activity (Han, Kim, Jeong, & Cohen, 2017). Since the neuroimaging studies demonstrated that selfhood significantly moderated moral emotion and motivation, the experiments manipulated the perceived distance between presented moral exemplars and participants’ self-concept.

The experiments presented two different types of exemplary stories: attainable and unattainable moral stories. Given the significant positive interaction between self-related and moral functioning-related brain regions, as the presented moral stories are perceived to be closely associated with participants’ self-concept, the motivating effect of the stories would become greater (Walton, Cohen, Cwir, & Spencer, 2012); attainable stories (e.g., stories of peer exemplars) would more strongly promote motivation compared to unattainable stories (e.g., stories of historic figures), which seem distant from participants. In fact, previous social psychological intervention experiments focusing on non-moral motivation also reported that attainable stories better promoted motivation while unattainable stories might backfire (Cialdini, 1980; Lockwood & Kunda, 1997).

First, a lab experiment was conducted to examine the motivating effects of different types of moral stories among college students; it used participation in voluntary service activity as a proxy for moral motivation. This experiment was conducted among 54 Korean college students. They were randomly assigned to one of these three groups: attainable exemplar, unattainable
exemplar, and control group. Participants in the attainable exemplar group were presented with stories of college students who participated in voluntary service activities less than three hours per week. Those in the unattainable exemplar group were presented with stories of peers who engaged in service activities more than nine hours per week. These groups were requested to write a brief letter persuading their friends to participate in service activities while referring to the presented exemplary stories. Non-moral stories, such as sports news reports were presented to the control group.

Second, a classroom intervention experiment tested the same hypothesis among 8th graders after an eight-week intervention period. This classroom-level experiment was performed to apply the lab-level intervention to more realistic educational settings. A total of 187 Korean 8th graders participated in this experiment. They were randomly assigned to one of these three groups: peer exemplar, historic figure, and control groups. The interventions were conducted for eight weeks (1 hour / week). The peer exemplar group was requested to discuss moral virtues of peer moral exemplars, e.g., family members and friends. The historic figure group was asked to talk about moral behaviors of historic moral exemplars, e.g., Mother Teresa and Martin Luther King Jr. Students in the control group were not presented with any concrete moral exemplars.

In both experiments, pre-test and post-test (about two months later) voluntary service engagements were compared between conditions. The findings from these two experiments supported the hypotheses (see Fig. 3). Attainable and peer exemplars better promoted moral motivation compared to unattainable and historic moral exemplars; these findings are coherent with the neuroimaging experiments that contributed to the hypothesis setting and previous intervention studies.
Applying Evolutionary Modeling and Computer Simulation to Inform Educators and Policy Makers

In order to apply the intervention programs that were developed in a lab or classroom to wider contexts, such as districts, more long-term experiments with diverse experimental conditions should be performed. These additional experiments are necessary to determine the type and frequency of interventions. Since the application of intervention programs to larger-scale contexts, such as their application to curriculum, may result in significant long-term changes in students’ development (Cohen et al., 2009; Yeager & Walton, 2011), these experiments would be important to educators and policy makers. However, due to the lack of time and resource, it would be difficult to conduct multiple long-term, large-scale experiments in real educational settings (Brown, 1992; Yeager, Fong, Lee, & Espelage, 2015). Given this
limitation, the evolutionary modeling and computer simulation predicting long-term outcomes of intervention programs can contribute to the large-scale implementation of developed interventions. For instance, predictions based on relatively small scale data can inform educators and policy makers of what kind of interventions would most effectively produce desirable developmental outcomes in larger educational settings in the long term.

The evolutionary modeling using the Evolutionary Causal Matrices (ECM) can predict the future status of a certain system consisting of different types of individuals (Claidière, 2009). The ECM predict the future status at $t_0+n$ from the status change between $t_0$ and $t_0+1$ with iterative calculations; with $n$ iterations, the predicted status at $t_0+n$ can be calculated (Claidière, Scott-Phillips, & Sperber, 2014). In case of the moral educational intervention, we can set the $t_0$ status as the pre-test voluntary service engagement and $t_0+1$ as the post-test engagement. By performing iterative calculations, we can compare the effectiveness of interventions according to their types and application frequencies. As presented in Fig. 4, the attainable exemplar-applied intervention can better promote engagement. Its effect size declines as the frequency of application gets lower and the intervention should be performed at least once per every 18 months to produce a large effect (Han, Lee, & Soylu, 2016). We remark that the ECM-based prediction is useful at predicting future outcome sequences based on a simple stochastic model with a relatively small number of estimated parameters.
Figure 4. ECM-applied intervention outcome prediction. (a) Status change between $t_0$ and $t_0+1$ in each condition using ECM. (b) Effect size ($\eta^2$) by intervention frequency. Red line: threshold for large effect size ($\eta^2 = .26$). Blue line: threshold for medium effect size ($\eta^2 = .13$). Green line: threshold for small effect size ($\eta^2 = .02$).

When a large enough amount of training data is in hand, one may apply machine learning algorithms in order to develop a data-driven prediction model. Among various machine learning algorithms, artificial neural networks with many layers, or simply “deep learning”, are the most popular choice of today due to its outstanding performance in many classical applications such as image classification, object recognition, speech recognition, etc. The deep architecture of deep learning corresponds to a hierarchy of features, factors, or concepts, where higher-level concepts are defined from lower-level ones, and the same lower-level concepts can help to define many higher-level concepts (Deng & Yu, 2014, p. 200). Using Google’s TensorFlow (Abadi et al., 2015), we trained a two-layered convolutional network for a predicting intervention outcomes.
More precisely, our prediction network takes pre-test variables (i.e., service engagement, gender, intervention type, emotional responses to intervention activity, intention to engage in service) as inputs, and predicts post-test outcome indicator (i.e., whether or not engage in service at the post-test). We used an iterative training algorithm (called the stochastic gradient method), and we observed that the prediction performance is maximized after about 4000 iterations of the training algorithm. The best prediction model with our convolutional network clearly outperforms simple logistic regression: while the accuracy of logistic regression was 75.47%, that of our best convolutional network reached 85.16%. These computer simulation models can provide researchers, educators and policy makers with a reliable way to predesign large-scale intervention experiments or applications.

1 Note that the prediction performance decreases after a certain number of iterations. This is called overfitting, which happens when a prediction model starts capturing in the model noise of the data, losing predictability.
Concluding Remarks

If and how findings from neuroscience studies can be beneficial for educational practice and policy making has been hotly debated in recent years (Ansari & Coch, 2006; Hruby, 2012; Immordino-Yang et al., 2007). Following the outline provided above we encourage the field to triangulate the information provided by lab studies, classroom interventions, computational modeling and meta-analyses of neural mechanism studies, to develop informed interventions based on empirically supported practices and procedures. Additionally, we see our approach using multiple empirical criteria for determining successful moral education interventions as providing a foundation for more general educational policy about what works in the moral domain and why. Taking these ideas a step further, we see our multi-source, multi-disciplinary approach as a blueprint for educational practice beyond the moral domain because it explicitly
heeds the call to build a two-way bridge between the lab and the classroom, and between neuroscience and education (Mason, 2009; Turner, 2011).

References


*TensorFlow: Large-scale machine learning on heterogeneous distributed systems.*

Mountain View, CA.


Sevinc, G., & Spreng, R. N. (2014). Contextual and perceptual brain processes underlying moral


