Considering the Purposes of Moral Education with Evidence in Neuroscience: Emphasis on Habituation of Virtues and Cultivation of Phronesis

Hyemin Han [1]

[1] Educational Psychology Program, University of Alabama

Author Note

Hyemin Han https://orcid.org/0000-0001-7181-2565

Correspondence concerning this manuscript should be addressed to Hyemin Han, Box 870231, University of Alabama, Tuscaloosa AL 35487. Email: hyemin.han@ua.edu

Abstract

In this paper, findings from research in neuroscience of morality will be overviewed to consider the purposes of moral education. Particularly, I will focus on two main themes in neuroscience, novel neuroimaging and experimental investigations, and Bayesian learning mechanism. First, I will examine how neuroimaging and experimental studies contributed to our understanding of psychological mechanisms associated with moral functioning while addressing methodological concerns. Second, Bayesian learning mechanism will be introduced to acquire insights about how moral learning occurs in human brains. Based on the overviewed neuroscientific research on morality, I will examine how evidence can support the model of moral education proposed by virtue ethics, Neo-Aristotelian moral philosophy in particular. Particularly, two main aims of virtue ethics-based moral education, habituation of virtues and cultivation of phronesis, will be discussed as the important purposes of moral education based on neuroscientific evidence.

Keywords: Moral education, Virtue ethics, Neuroscience, Habituation, Phronesis

Competing Interests

The authors did not receive support from any organization for the submitted work. The authors have no relevant financial or non-financial interests to disclose.

Acknowledgments

The author thanks Peter Königs, Gregor Hochstetter, Hanno Sauer, Paul Rehren, Kirsten Meyer, Dominik Balg, Alex Madva, and other workshop members as well as three anonymous reviewers for their constructive comments on an earlier version of the manuscript. The author also appreciate Kelsie J Dawson's feedback on the revised manuscript.

Considering the Purposes of Moral Education with Evidence in Neuroscience: Emphasis on Habituation of Virtues and Cultivation of Phronesis

Abstract

In this paper, findings from research in neuroscience of morality will be overviewed to consider the purposes of moral education. Particularly, I will focus on two main themes in neuroscience, novel neuroimaging and experimental investigations, and Bayesian learning mechanism. First, I will examine how neuroimaging and experimental studies contributed to our understanding of psychological mechanisms associated with moral functioning while addressing methodological concerns. Second, Bayesian learning mechanism will be introduced to acquire insights about how moral learning occurs in human brains. Based on the overviewed neuroscientific research on morality, I will examine how evidence can support the model of moral education proposed by virtue ethics, Neo-Aristotelian moral philosophy in particular. Particularly, two main aims of virtue ethics-based moral education, habituation of virtues and cultivation of phronesis, will be discussed as the important purposes of moral education based on neuroscientific evidence.

Keywords: Moral education, Virtue ethics, Neuroscience, Habituation, Phronesis

Introduction

In this paper, I will overview findings from recent neuroscientific research on morality to consider what the purposes of moral education shall be. While discussing the topic, I plan to refer to moral philosophical accounts regarding moral education, instead of solely focusing on neuroscience. Research in neuroscience might be able to provide moral educators with useful insights about psychological processes involving moral functioning and development at the biological level (Han 2016). However, building any normative or prescriptive argument about what shall be pursued in moral education solely based upon empirical evidence would not be convincing. This is because such evidence would not necessarily provide useful ideas about value-related consideration, what should be pursued and taught during the course of moral education, in a direct manner (Kristjánsson 2013; Bowers 2016). Hence, my consideration of the purposes of moral education with neuroscientific evidence will be informed by moral philosophy, particularly Neo-Aristotelian virtue ethics, which is primarily interested in habituation of virtues and cultivation of practical wisdom (Sanderse 2014).

The main purpose of this paper is not to support one moral philosophical position, virtue ethics, against others at the normative level. Instead, I intend to discuss how suggestions on moral education proposed by virtue ethics, Neo-Aristotelian in particular, are supported by evidence regarding moral psychology and development, and further, how the interventions and educational methods based on them are more capable of producing the most effective educational outcomes compared with others. We can consider why examining different educational standpoints based on scientific evidence practically makes sense (Han 2016). For example, in physical education, findings from physiology and anatomy informed how to design effective educational programs. Instructions in physical education were thus developed to be consistent with such scientific evidence (Bulger et al. 2008). In the same vein, moral education can also be designed with scientific evidence so that it can more effectively promote moral development.

To achieve this purpose, first, I will review recent findings from large-scale data-driven explorations focusing on morality at the neural level. Meanwhile, I intend to discuss limitations in conventional neuroimaging studies and how they could be alleviated through novel analysis methods in recent neuroscience research (Yarkoni et al. 2011; Dockès et al. 2020). Furthermore, I plan to overview experimental studies that employed diverse investigational methods that allow researchers to examine the causal relationship between the neural activity and psychological process of interest (Bestmann and Feredoes 2013). Such studies would provide useful ideas about neural activity patterns associated with moral functioning and how learning, in addition to motivational and behavioral changes, occur through neuroscience-informed exercises and trainings. Based upon the review of large-scale neuroimaging studies and experimental studies, I will discuss which psychological processes shall be focused on with conceptual consideration on Neo-Aristotelian virtue ethics.

In addition to neuroimaging and experimental studies, I will also overview recent discussions on the Bayesian nature of learning processes in human brains (O'Reilly et al. 2012). By overviewing the Bayesian perspective, I will consider the neural-level mechanism of learning based on the neural network training (Churchland 1998), which can also be applicable to moral learning (Cushman et al. 2017), and how such learning can be promoted to produce optimal outcomes in the long run. I intend to connect the Bayesian nature of moral learning at the neural level and the ideas for moral education proposed by Neo-Aristotelian virtue ethics to explore the purpose of moral education.

Evidence from Neuroimaging Analyses and Experimental Studies

Examining moral functioning at the neural level may facilitate our understanding of the psychological substrates of moral functioning. It may also inspire a discussion on the purposes of moral education by providing ideas about psychological processes constituting moral functioning (Han 2016; Han et al. 2019). Although the account supporting the value of neuroscience in research on morality might be intriguing, there have been continuous debates about whether neuroscientific studies can provide useful insights about the basis of psychological functioning of interest. For instance, researchers have been concerned about whether we can clearly identify the pattern of neural activity specifically associated with one cognitive function of interest with poor generalizability due to methodological limitations (Eickhoff et al. 2020). Also, some have criticized that the majority of conventional neuroimaging studies can only illuminate correlation, not causality (Ramsey et al. 2010). Thus, before discussing how neuroscience can potentially contribute to moral education with information about the psychological mechanism of moral functioning, I intend to overview these concerns and discuss how recent methodological advances can alleviate them. Then, I will examine neuroscientific studies conducted with the advanced methods.

Methodological benefits of large-scale neuroimaging analyses

First, methodological limitations existing in conventional neuroimaging techniques prevent us from directly matching a specific functionality of interest to a specific region showing significant activity (Eickhoff et al. 2020). With conventional neuroimaging findings, which demonstrate the differences in brain activity in different task conditions, we cannot fully examine whether such activity is particularly associated with the task condition of interest. Many fMRI experiments examining moral functioning compare neural activity between two specific task conditions, e.g., Trolley dilemma vs. non-moral arithmetic decision making task conditions (Greene et al. 2001), with a small sample size resulting in weak statistical power (Eickhoff et al. 2020). Hence, even if we can identify patterns of significantly activated brain regions from such experiments, it is not convincing to argue that such a pattern reveals a unique pattern of neural activity associated with a functionality of interest.

Fortunately, several novel analysis tools seem capable of addressing this methodological issue in neuroimaging analysis at least partially with information technology (Poldrack and

Yarkoni 2016). NeuroSynth and NeuroQuery use a large-scale neuroimaging database to address the issue (see supplementary materials for further details about the tools). Statistical maps created by such tools demonstrate the unique neural activity pattern associated with a specific functionality by analyzing a large-scale database containing fMRI experimental results addressing various functionality topics (Rubin et al. 2017).

Let us examine the brain activity pattern uniquely associated with morality, the main topic of this paper. Both NeuroSynth and NeuroQuery reported significant activation in the default mode network (DMN), which is associated with social cognition and self-related processes (Han 2017; Koban et al. 2021), when a keyword "moral" was entered (see the "methodological further details" section in the supplementary materials for methods and the "results from large-scale neuroimaging analyses" section for the full results). In addition, I also employed the NeuroSynth decoder to explore the neural activity patterns most significantly and uniquely associated with moral functioning (Yarkoni et al. 2011). This reported that keywords related to theory of mind (or mentalizing) and self-referencing (or autobiography) showed the most significant association. These results may suggest that the neural circuitries and psychological processes related to morality and moral functioning are uniquely associated with theory of mind and self-related processes.

Let me consider why this decoding result can provide novel insights that could not be gained by conventional neuroimaging experiments. As an illustrative example, let us imagine a case of decoding the fundamental components of decathlon, which was employed by Kristjánsson and Fowers (2022) to examine the nature of character education. If we use an experiment method that is like conventional neuroimaging, then at the best, we might be able to compare the mechanisms between decathlon[i] vs. control (e.g., other sports) conditions. Even if we discover that 100m run and pole vault mechanisms more frequently appeared in the decathlon condition compared with the control condition, such a result may provide very limited evidence supporting the unique pattern of the mechanism of decathlon due to the peculiarity of the employed conditions and statistical power. However, once we enter the data acquired from decathlon into a decoder like the NeuroSynth decoder, which enables us to examine a large-scale dataset containing mechanisms of diverse sports, then it would be able to show us the unique mechanism of decathlon. Once we see that the mechanisms of the 100m run and pole vault report the highest correlation with decathlon, then at the least, we can better conclude that these two are uniquely associated with decathlon.

One point to note is that the uniquely associated features suggested by the NeuroSynth decoder can provide information about which features may be considered as lower-level features of the functionality of interest (Rubin et al. 2017). Given what has been discussed in the field, particularly related to the hierarchical structure of cognition, the concrete and specific components at the lower level identified through the decoding process can be considered as the components being employed for the more complex psychological functionality (Schurz et al. 2021). Both psychologists and philosophers suggest that moral functioning is very complicated and such abstract cognitive functioning employs diverse concrete cognitive functionalities, including theory of mind and self-related processes (Colby and Damon 1993; Blair 2007; Young et al. 2010). Therefore, with the NeuroSynth decoder based on a large-scale dataset (Klein 2012), we can better identify concrete lower-level cognitive functionalities that are uniquely associated with and employed for moral functioning (Poldrack and Yarkoni 2016).

Of course, the currently introduced methods are not perfect. Because NeuroSynth, NeuroQuery, and NeuroSynth decoder use keywords, which have not been carefully filtered, the set of neuroimaging data employed do not necessarily and exclusively be related to morality as defined by theories (Yarkoni et al. 2011). Also, because they are still correlational, the results do not necessarily show us any causal association. Despite the limitations, the novel methods are more capable of identifying the unique neural-level patterns of moral functioning thanks to the large-scale database and sophisticated data-driven analysis compared with individual neuroimaging studies focusing on specific task conditions. Such findings will be able to provide us with insights about the unique functional features of morality that could not be obtained with traditional methods.

Issue of functional localization and network-based approaches

Second, neuroscientists argue that a psychological functionality of interest could not be exclusively associated with only one specific region showing significant activity (Ramsey et al. 2010). Thus, we need to explore neural networks or circuitries given recent findings from neuroimaging studies employing network-based approaches (Bressler and Menon 2010; Ramsey et al. 2010). Network-based approaches in neuroimaging, such as the psychophysiological interaction analysis and connectivity analysis methods (Friston et al. 1997; Cáceda et al. 2011), have enabled researchers to examine brain circuitries associated with psychological functions of interest quantitatively.

Recent research on the neuroscience of morality has also reported brain circuitries associated with morality (see the "results from studies employing network-based approaches" section in the supplementary materials for more information). Consistent with the large-scale neuroimaging analysis, significant connectivity between the DMN regions and other brain regions associated with cognition, emotion, and motivation within the moral task contexts has been demonstrated in previous studies (Cáceda et al. 2011; Pujol et al. 2012; Han et al. 2016).

Although there have been several debates about whether the DMN is functionally associated with self-related psychological processes, additional evidence from recent neuroimaging studies conducted have supported the point. In the previous studies conducted with both human and non-human subjects, DMN activity significantly decreased when the subjects were sleeping or anesthetized compared with when they were awake (e.g., Sämann et al. 2011; Li and Zhang 2018). Recent research, including both from ordinary and clinical populations, has also consistently suggested that the DMN plays fundamental roles in self-related cognition within social contexts instead of mere idling (for review, Koban et al. 2021).

Issue of causality and experimental methods in neuroscience

Furthermore, we also need to consider concerns regarding the correlational nature of conventional neuroimaging methods. Given correlation per se does not necessarily prove causality, even if we find significant association between psychological functionality A and activity in brain region B, we cannot confidently argue that A is a result of activity in B even with enhanced neuroimaging methods (Ramsey et al. 2010; Bestmann and Feredoes 2013). Hence, it would be necessary to consider how to address the issue of causality before discussing how neuroscience can inform our exploration on the purpose of moral education, which involves norms and prescriptions in addition to descriptions.

Several experimental methods in neuroscience at least partially enable us to examine causality between activity in specific brain regions or networks and psychological functionalities of interest. Such methods include neurostimulation, neurofeedback, and neuroscience-informed intervention (Bestmann and Feredoes 2013). Given education, including moral education, is primarily concerned with how to promote one’s development and growth through educational activities (Yeager and Walton 2011), it would be important to address the causality issue if we intend to learn from neuroscience for improving education. Thus, overviewing findings from such experimental studies may provide us with useful insights about mechanisms associated with learning and development, which are fundamental in educational research (see the "results from experimental studies" section in the supplementary materials for more information).

First, neurostimulation employing the Transcranial Magnetic Stimulation (tMS) and Transcranial Direct Current Simulation (tDCS) enabled researchers to examine the causal relationship between a functionality of interest and a brain region or circuitry by temporarily altering activity in the brain region or circuitry of interest (Young et al. 2010; Bestmann and Feredoes 2013; Riva et al. 2019). Previous neurostimulation studies have demonstrated that influencing regions and circuitries associated with moral functioning, e.g., the DMN, significantly changed moral cognition and behavior, e.g., evaluating self-traits, evaluating others' intent regarding moral transgression, moral decision making, and compliant behavior (Lou et al. 2010; Young et al. 2010; Ruff et al. 2013; Riva et al. 2019).

Second, neurofeedback studies, which apply real-time neuroimaging techniques to provide participants with information about how to regulate targeted psychological processes, can also provide additional evidence supporting causality (Fede et al. 2020; see supplmentary materials for methodological further details). This method has been reported to produce neural activity in regions of interest during intervention sessions as well as long-term improvement in targeted functionalities (Hohenfeld et al. 2017). For instance, applying neurofeedback to monitor neural activity in the DMN during psychological interventions significantly contributed to improving mentalizing among clinical populations at both neural and behavioral levels (Imperatori et al. 2017). Thus, neurofeedback would be a possible way to support a causal relation between brain regions or circuitries and psychological functionalities.

Furthermore, evidence from neuroscience-informed intervention studies aimed at promoting specific behavioral outcomes, such as moral behavior, can also be considered (e.g., Han et al. 2017). Han et al. (2017) reported that relatable close-other moral exemplars, such as friends, better motivated prosociality compared with unrelatable distant exemplars, such as historic figures (see also Han et al. 2022). They assumed that the perceived distance and connectivity between the presented exemplars and participants would be a significant factor determining psychological impacts of exemplars based on Han et al. (2016), which reported significant involvement of the DMN in moral functioning. If such neuroscience-informed interventions work as intended, the fact may at least partially support a causal relation between brain activity and psychological functioning. It may also suggest that neuroscience can inform education, including moral education.

Moreover, trainings and interventions aimed at behavioral changes have also been known to induce changes in structure and functioning at the neural level (Johansen-Berg 2012). This may support the point that brain activity and psychological or behavioral functioning may causally and reciprocally influence each other. Across diverse domains of affective, cognitive, and motivational processes, including those closely related to moral functioning such as compassion and perspective taking, psychological training programs have significantly affected the brain structure, connectivity, and functioning in relevant brain regions (Lampit et al. 2015; Valk et al. 2017; Tymofiyeva and Gaschler 2021). Such evidence unequivocally suggests that well-designed long-term trainings can induce changes at the motivational, behavioral, and even neural levels.

Taken together, with novel methods such as neurostimulation, neurofeedback, and neuroscience-informed intervention, neuroscience research allows us to examine the causal relations between brain regions and circuitries, and psychological functioning of interest. Researchers can now better understand how activity in certain brain regions or circuitries influences and regulates psychological and behavioral processes. Such understanding would eventually widen our knowledge about the neural- and biological-level substrates of human psychology and behavior.

Implications of Evidence from Neuroimaging Analyses and Experimental Studies on Virtue Ethics-based Moral Education

In this section, I intend to discuss how neuroscientific evidence can inform exploration of the purpose of moral education. First, in terms of developmental processes, neuroscientific studies employing advanced techniques, particularly those employing neurofeedback or interventions, may suggest that well-designed trainings can induce long-term functional as well as structural changes in brain regions associated with the targeted functioning (Hohenfeld et al. 2017; Tymofiyeva and Gaschler 2021). Furthermore, changes occur not only at the neural level but also at the psychological and behavioral levels in the long term (Valk et al. 2017).

Experimental studies demonstrated that interventions can be developed in a way where they can pinpoint psychological processes of interest with neuroscientific evidence. Such a point can also be applicable to the domain of moral functioning. It may shed light on our exploration of what shall be focused on in moral education. If we can identify brain activity patterns associated with moral functioning, interventions in moral education may aim at regulating activity in the associated brain regions (Han 2016; Han et al. 2017). Then, with appropriate intervention approaches involving long-term trainings, it would be possible to induce behavioral and neural-level changes in the targeted moral functioning and its associated brain structures in the long run.

Let us revisit the example of the decathlon training to examine why aiming at functionalities suggested by neuroimaging evidence shall be seriously considered in moral education. According to Kristjánsson and Fowers (2022), balanced training in each individual sport constituting the decathlon is required for optimal performance. Although mastering the individual sports is not sufficient, it is necessary to train skills in such individual sports at the least. Similarly, in the moral domain, if we can identify functional components for moral functioning, e.g., self-related processes and mentalizing to understand others identified by neuroimaging, becoming well-equipped with such components may be considered a necessary condition for optimal moral functioning. In fact, Kristjánsson and Fowers (2022) also argued that to become a fully virtuous being "presupposes habituation into individual moral virtues (p. 12)". Thus, like the case of a decathlon, it would be plausible to say that developing the aforementioned identified functional components is a necessary condition for optimal moral functioning, which shall be pursued in moral education.

The aforementioned point may support one of the aims of moral education proposed by virtue ethics, habituation of virtues through long-term exercises, with evidence at the neural level (Steutel and Spiecker 2004; Sanderse 2014). As shown, behavioral trainings resulted in substantial changes in brain regions and circuitries. Functional and structural changes in brains induced further behavioral changes in the long term. These findings suggest that well-organized trainings and exercises can induce intended changes in the targeted psychological functions at the neural level (Lampit et al. 2015; Tymofiyeva and Gaschler 2021), and eventually reinforce behavioral changes (Valk et al. 2017). Although directly relevant neural-level evidence has not been sufficiently accumulated in the domain of morality, such a point might also be applicable to moral education. Virtue habituation through repeated exercise, particularly that happens during childhood (Sanderse 2020), when brains are highly plastic (Tymofiyeva and Gaschler 2021), is perhaps capable of promoting neural-level changes in brain regions and circuitries associated with moral functioning in the long term. Then, as Neo-Aristotelian moral philosophers have argued, habituated virtues, which might have associated neural-level mechanisms, may lead to intended behavioral outcomes (Sanderse 2020). Consequently, neural-level evidence related to the reciprocal relationship between neural-level and behavioral-level changes reported by neurofeedback and intervention studies may suggest that habituation of virtues shall be one of the main purposes of moral education as virtue ethicists proposed.

Of course, some may argue that the mechanism of habituation presented in neuroscience is not necessarily identical to habituation of virtues, but more likely to be mere habituation of behavior. Although many neuroscientific experiments have employed relatively simple experimental paradigms, some studies can provide information about the mechanism of virtue habituation. Recent neurofeedback research has demonstrated that higher-order social cognition, including theory of mind, self-related processes, and perspective taking, which are fundamental in moral functioning (Kristjánsson et al. 2021), can be promoted at the neural-, psychological-, and behavioral-levels (Imperatori et al. 2017; Valk et al. 2017). Also, from the philosophical perspective, Navarini (2020) proposed that virtue habituation shall be implemented in terms of both neural- and behavioral-level changes through learning and experience over time. In fact, Churchland (1998) also argued that acquisition of moral virtues through moral learning resulting in habituation shall be substantiated by synaptic readjustments at the lower level, and such readjustments are supposed to constitute changes at the higher level. Given these neuroscientific as well as philosophical accounts, we may assume that virtue habituation involves functional and structural changes in brains, particularly in regions associated with moral functioning, in addition to long-term behavioral changes.

Second, the other important aim of moral education in virtue ethics, cultivation of phronesis (Darnell et al. 2019), can be supported by evidence from recent neuroimaging studies as well. The results from the large-scale neuroimaging data analyses have unequivocally reported significant association between moral functioning and the DMN (Han 2017). Although the DMN involves multiple psychological functions (Buckner et al. 2008), the results from the NeuroSynth decoder may provide more specific information about which psychological functions are uniquely correlated to moral functioning. The NeuroSynth decoder demonstrated that two main themes, theory of mind and self-related processes, were most uniquely associated with moral functioning. Theory of mind is about understanding and inferring others' mental states and beliefs properly (Young et al. 2010). Within the context of moral functioning, self-related processes are about referencing one's beliefs and values. Such processes are also related with reflection and deliberation upon existing beliefs (Colby and Damon 1993; Han 2017).

The large-scale neuroimaging evidence may suggest that moral functioning is inseparable from considerations and deliberations on others (the theory of mind) and oneself (self-related processes) (Colby and Damon 1993; Young et al. 2010). Furthermore, findings from brain stimulation studies that disrupted brain regions associated with self-related processes or mentalizing reported significant changes in moral decision-making and behavior, which also supports the point (Young et al. 2010; Ruff et al. 2013; Riva et al. 2019). Given these two functionalities were most strongly associated with moral functioning consistently across neuroimaging and experimental research, we may assume that these are central and even fundamental in moral functioning in general.

From the Neo-Aristotelian perspective, the two functionalities, self-related processes and mentalizing, can be considered important constituents of phronesis, which is required to render the most appropriate decision (Kristjansson 2015), and ultimately, optimal moral functioning (Sanderse 2014, 2020). According to Kristjansson (2015), phronesis is a virtue to "'deliberate finely' about the relative weight of competing values, actions and emotions in the context of the question of 'what promotes living well in general' (p. 88)." For optimal decision making, one needs to perceive and evaluate external situational factors, including environmental factors as well as others' intents and thoughts. This can only be done with well-developed mentalizing abilities (Bzdok et al. 2012; Kristjánsson et al. 2021). Furthermore, one should also be able to understand and evaluate one's own beliefs and values well (Kristjánsson 2005; Han et al. 2017). Hence, abilities to understand and evaluate oneself (Noel 1999; Kristjansson 2015) and others (Kristjansson 2015; Lapsley 2021) are deemed to be fundamentally required for phronesis (Kristjansson 2015). Consequently, the findings from the large-scale neuroimaging analyses and experimental studies demonstrating the unique tie between the aforementioned two functionalities and moral functioning would support the importance of phronesis in moral functioning in general.

However, there may be a concern that the neuroscience studies cannot shed light on virtue ethics because they examined ordinary people, not virtuous beings. Although I admit the point, alternatively, we may refer to several studies that compared the activity and anatomy in the brain regions of interest, such as the DMN, between different populations. We may particularly refer to research on psychopathy and human development. It would be plausible to assume that psychopaths who actively engage in antisocial behavior without affective and motivational capacities in the moral domain are less virtuous than ordinary people in general

(Vance and Werner 2022). Also, we may note an argument that virtue development is gradual (Athanassoulis 2000). If so, then, comparing the brains between psychopaths and ordinary people, or even within ordinary people at different levels of moral development may provide useful information.

Previous studies have reported that psychopaths demonstrated significantly lower activity and more structural deficiencies in the regions associated with moral functioning, e.g., the DMN, compared with ordinary populations (Raine and Yang 2006; Lenzen et al. 2021). Moreover, among ordinary populations, participants with more sophisticated moral reasoning, in terms of post-conventional reasoning, showed higher gray matter thickness in the regions (Prehn et al. 2015). Thus, these results may inform the neural-level aspects of "more virtuous beings" at least indirectly.

Bayesian Learning Rule and Moral Learning

Recent research in computational neuroscience has proposed that learning in brains follows the Bayesian rule (O'Reilly et al. 2012; Cushman et al. 2017; see supplementary materials for further details about the rule). Various learning models proposed in computational neuroscience, particularly the reinforcement learning model, which is frequently applied to the modeling of learning social cognition including moral cognition, are commonly based on the Bayesian rule (Hackel and Amodio 2018). Through the Bayesian learning process, we adjust our beliefs, which are used to predict what would happen in the future. In an ideal situation, well-adjusted beliefs are likely to minimize prediction errors, the differences between predictions and actual outcomes, and to help us prepare future situations more effectively and quickly. If what actually happens contradicts our prior beliefs, and thus our prediction, then we adjust the beliefs based on the prediction error (Kim et al. 2020). However, adjusting our beliefs and improving prediction accuracy do not necessarily minimize the future errors, and thus, produce ideal outcomes in all instances. One major concern is overfitting (Babyak 2004). Overfitting occurs when our beliefs are overly adjusted by inputs from a very specific domain. In such a situation, the overfitted prediction model may accurately predict outcomes within the specific domain; however, prediction accuracy significantly worsens when predicting novel situations outside of the domain (Han et al. 2020).

Although overfitting is a significant problem in learning, we may employ regularization during the learning process to address it (Zou and Hastie 2005). To understand this mechanism better, how regularization occurs in deep learning will be overviewed. While conducting the deep learning, which simulates the neural network system to construct a prediction system in a computational manner, regularization is commonly employed to prevent overfitting (Han et al. 2020). It is used to adjust the extent to which inputs affect the adjustment of the weights of neurons, which are related to the belief-updating mentioned previously. This mechanism of the neural network system is also applicable to moral learning as well. Churchland (1998) referred to the neural network model to explain how learning of moral knowledge and virtue occurs at the neural level. According to his explanation, one acquires moral knowledge and virtue by training a neural network through observations, reinforcements, and weight adjustments.

Without regularization, inputs directly determine weights, and eventually, updated outputs. In such a case, inputs from a limited domain completely determine the weights, so the whole neural network becomes overfitted to such limited inputs and cannot predict incidences outside of the domain accurately (Han et al. 2020). The mechanism of regularization alleviates such an issue by downregulating the influences of the inputs on weight adjustment. While training a neural network, one can determine regularization parameters, which regulate to what extent inputs contribute to weight adjustment. The determined regularization parameters influence the overall performance of a trained neural network, prediction accuracy as well as robustness against overfitting (Epskamp and Fried 2018). Hence, how to determine the regularization parameters becomes an important matter in the learning process. Once the regularization parameters are properly determined, the trained neural network should be able to predict outcomes accurately within as well as outside of the domain of provided inputs, as it is not overfitted to the inputs.

Similarly, in the reality, humans also employ regularization when they adjust their prior beliefs while encountering situations that challenge and contradict the beliefs. For instance, in a recent empirical study, psychologists examined how scholars change their prior beliefs about effects of experimental manipulations in psychological studies after being presented with evidence challenging the beliefs from replication studies (McDiarmid et al. 2021). They predicted to what extent the researchers were likely to change their prior beliefs with Bayes theorem. The observed adjustment of prior beliefs was significantly smaller than what was expected by Bayes theorem, although the prior beliefs were updated somehow. It may suggest that beliefs are not directly and completely influenced by external inputs, so the influences from such inputs perhaps are attenuated through regularization.

Implications of Learning Mechanism based on Bayes Rule on Virtue Ethics-based Moral Education

The aforementioned Bayesian learning rule may provide additional support to the Neo-Aristotelian model of moral development at the neural level given its conceptual similarity to the Neo-Aristotelian model. Thus, it may provide insights about what shall be pursued by moral education as well. Now, I will focus on two aspects in the Bayesian learning rule, which has significant connectivity with the Neo-Aristotelian model. First, setting prior beliefs is associated with habituation in virtue ethics. Second, the importance of phronesis can be considered within the context of regularization and its parameter determination.

First, the importance of prior formulation in the Bayesian learning mechanism may suggest the importance of early habituation and internalization of desirable moral virtues and values proposed in Neo-Aristotelian moral education. Within the Bayesian framework, learning occurs in terms of Bayesian updating of priors, prior beliefs that one possesses before observing data. Through data observation, prior beliefs are updated following Bayes theorem and then eventually become posterior beliefs. Although observed data updates the prior beliefs, the prior beliefs per se also significantly influence the posterior beliefs. For instance, if prior beliefs are strongly biased to direction A, then observing data that promotes updating toward direction B is less likely to change the prior beliefs compared with when the prior beliefs were less biased. In fact, when Bayesian inference is conducted, the outcome is sensitive to prior beliefs (Sinharay and Stern 2002). Also, use of appropriate priors in Bayesian inference is a way to minimize a potential bias (Han 2022). Such a point suggests that setting up prior beliefs is very important to produce desirable outcomes at the end of the Bayesian learning process.

Such a point can be applicable to moral education if the formation of one's moral beliefs and perspectives also follows the Bayesian learning mechanism (Kim et al. 2020). Let us assume that experiences gained through interactions with external circumstances (e.g., peer interactions, educational activities) constitute observed data for Bayesian learning (Cushman et al. 2017). Then, the prior beliefs correspond to prior moral beliefs and views that one previously possessed. If the quality of priors (e.g., biasedness) significantly influences the posterior beliefs, then within the moral domain, possessing morally appropriate prior beliefs would become important for optimal developmental outcomes. From the Neo-Aristotelian perspective, it can be explained in terms of successful habituation and internalization of desirable moral virtues and values (Sanderse 2020). Regarding this point, Navarini (2020) suggested that neuronal tracing occurs when one involves continuous virtue learning and then shapes internalized virtuous habits. Such neuronal tracing promotes the likelihood of virtuous action at the end.

If one successfully habituates and internalizes moral virtues during their childhood, then they become optimal priors, which contribute to the eventual formation of optimal posteriors (Kristjánsson 2013). On the other hand, if one is exposed to and internalizes extreme and biased values and views, then they are likely to constitute suboptimal prior moral beliefs. Although long-term updating of the priors may lead to less biased posteriors, compared with the case when the ideal priors were formed, it would be difficult to promote optimal developmental outcomes with presence of biased priors (Sinharay and Stern 2002). As early prior formulation significantly influences eventual posterior outcomes, early habituation of virtues may also significantly determine the eventual moral developmental outcome.

Furthermore, the regularization mechanism in Bayesian learning may support the importance of cultivation of phronesis as an aim of moral education. The neural network example illustrated that regularization parameters should be carefully determined as they are responsible for the extent that external inputs influence the adjustment of neural weights, which significantly determine posterior beliefs. Determining regularization parameters is critical in minimizing both under-fitting and over-fitting of the prediction model (Zou and Hastie 2005; Epskamp and Fried 2018). This mechanism is also applicable to human learning processes (McDiarmid et al. 2021). Posterior moral beliefs are not merely updated from their prior beliefs with external inputs, such as moral lessons, interactions with close others, etc. If we completely update our moral beliefs with external inputs, then our moral system is likely to be overfitted with inputs from neighboring environmental factors. Of course, our moral system can address familiar situations intuitively and cost-effectively if the system was well trained with familiar inputs (Greene 2014). However, as Greene (2014) argued, if our judgment system is overfitted with a specific set of values and we tend to make judgments based on such values, then we are likely to render inappropriate decisions when we are dealing with unfamiliar situations, particularly when we encounter conflicting diverse values and perspectives. Thus, in moral learning, regularization is required to prevent such an overfitting issue within the moral domain.

The regularization process during moral learning requires very sophisticated parameter adjustment for optimal learning outcomes. Let us consider some illustrative examples. On the one hand, in some instances, we may need to update our prior beliefs and prediction system in the moral domain radically. Let us imagine ordinary people in a totalitarian country encountering radical but morally justifiable arguments, such as a call for human rights. From their perspectives, the aforementioned input, a call for human rights, is perceived to contradict their prior moral beliefs radically. However, if we consider what the morally appropriate action to take in such a situation is, they shall adjust their prior beliefs, and thus, also significantly update prediction systems affecting moral motivation and behavior based on the input. On the other hand, in some cases, we shall not update our beliefs and prediction system based upon contradicting inputs. For instance, an appropriate way to deal with extremist views is one example. Although the extremist views have become prevalent in modern society, if we adjust our prior beliefs with such inputs directly, it may end up with an increase in extremist behavioral tendency eventually. Instead, impacts from such extremist views shall be attenuated through the regularization mechanism during the moral learning process.

These two illustrative examples suggest that: first, our moral beliefs are not and should not be merely and passively updated by external inputs; and second, we need to determine to what extent our posteriors are updated with such inputs depending on contextual factors. If our moral system is directly and completely altered by external inputs as is a neural network without regularization, then we are likely to make inappropriate decisions particularly when we encounter unfamiliar situations due to overfitting. These two points substantiate the importance of determining and tuning the regularization parameters in moral learning processes for optimal moral learning outcomes.

One difficult point is that we need to determine our regularization parameters in moral learning very carefully across different situations as shown in the two illustrative examples. That says, for optimal moral learning, one must be capable of finding out the most appropriate regularization parameters depending on situational and contextual information. Consequently, one should possess sophisticated practical wisdom to be able to do so as proposed in virtue ethics (Kristjánsson et al. 2021). Given ourselves, not others, should determine the parameters, we need to possess well-developed practical wisdom. Because optimal moral learning could not be achieved without sophisticated practical wisdom to identify the optimal regularization parameters, cultivation of phronesis, which is underscored by virtue ethicists, should be a central aim of moral education. In fact, this point is consistent with De Caro et al.'s (2021) argument that phronesis is a sort of ethical expertise formed through education and training. Through cultivation of phronesis, we can optimize moral learning across different situations by appropriately adjusting regularization. With such optimized regularization and moral learning, eventually, we can perceive external inputs, and then, update our beliefs and moral system appropriately.

Although the weight adjustment mechanism may provide useful insights to understand the theme of phronesis cultivation better, we also need to consider an issue related to this matter. As shown by McDiarmid et al. (2021), it is practically difficult to induce radical posterior adjustments when such adjustments are deemed to be appropriate and needed (e.g., promoting views supporting human rights in a totalitarian society). Then, although cultivation of phronesis for appropriate weight readjustment may be the ideal solution in this case, for ordinary students, who are unlikely to possess sophisticated phronesis, doing so is very challenging in fact. To address this issue, we may refer to education based on communities and exemplars, which is also supported by virtue ethics (Kristjansson 2015). If students can be frequently exposed to moral community members and exemplars, they can provide the students with moral inspirations in a continuous manner (Curren 1999), so such continuous stimulations will induce more vigorous posterior updates as well as eventual cultivation of phronesis (De Caro et al. 2018). In such a case, although a one-time radical adjustment would not occur, the posterior beliefs will eventually be updated in a virtuous way thanks to continuous stimulations and updates.

Consequently, the Bayesian learning mechanism proposed in neuroscience, which employs regularization for weight adjustment, may support another important theme of virtue ethics-based moral education, cultivation of phronesis. Without well-cultivated phronesis, neither rendering the most appropriate decision within a given situation nor optimal moral learning with proper regularization, which is required for moral development, could be achieved (Sanderse 2014; Kristjánsson et al. 2021).

Concluding remarks

In this paper, I overviewed neuroscientific studies of moral functioning to examine what shall be pursued in moral education with ideas from virtue ethics, Neo-Aristotelian in particular.

Two streams of neuroscientific studies have been overviewed: first, large-scale neuroimaging analyses and experimental studies examining the neural activity patterns of moral functioning; and second, a neurocomputational approach to moral learning based on the Bayesian perspective and regularization. Regarding neuroimaging and experimental studies, I proposed that novel methodologies can address limitations in conventional neuroimaging studies. Large-scale neuroimaging analyses and network-based approaches allow researchers to address issues related to mere localization of psychological functioning and to identify the unique neural-level mechanism of interest. Experimental studies, such as neurostimulation, neurofeedback, and intervention studies, can examine causal relations between brain regions or circuitries and psychological functioning of interest. Furthermore, the mechanism of Bayesian learning with regularization in brains can provide us with useful insights about how moral learning and development occur.

Accordingly, two main themes regarding moral education in virtue ethics, early habituation and cultivation of phronesis, have been considered to be supported by evidence in neuroscience. First, the importance of habituation and internalization of virtues in moral education can be supported by evidence from neuroscientific experiments and the Bayesian learning mechanism. Evidence from experimental studies suggest that habituation through well-designed continuous exercises will produce significant changes in both brain functioning and structures, and behavioral outcomes reciprocally. Moreover, the importance of prior beliefs in Bayesian learning underscored that establishing appropriate prior beliefs is fundamental to promote optimal learning. Early habituation of virtues and values are necessary to found one's prior beliefs upon morally appropriate values so that one can have desirable posterior beliefs in the long run.

Second, given evidence from large-scale neuroimaging analyses and the Bayesian learning mechanism with regularization, cultivation of phronesis shall also be emphasized as a goal of moral education. Findings from large-scale meta-analyses of neuroimaging studies of morality, including those from NeuroSynth and NeuroQuery, demonstrated that both self-related processes and mentalizing are fundamentally associated with moral functioning after controlling for the baseline brain activity across diverse task conditions. These two functionalities constitute the basis of practical wisdom to perceive and evaluate one's own beliefs and values as well as external factors, such as others' intent, and thus, to render the most appropriate decision within a situation. Furthermore, the regularization process in Bayesian learning suggests that cultivation of phronesis is also important for optimal moral learning. Well-developed phronesis is necessary in determining appropriate regularization parameters, which are required for optimal moral learning across diverse situations.

Although comparison between different philosophical viewpoints is not the central theme of this paper, let me briefly discuss why the overviewed evidence may best support virtue ethics, the Neo-Aristotelian perspective in particular, better than others. The most important point to note is that Neo-Aristotelian virtue ethics underscores cooperation between reasoning and emotion, and deliberation and habits, which is inseparable from the development of a moral agent rather than the mere promotion of explicit moral choice or action, in moral functioning and moral development (Anscome 1958; Jordan and Kristjánsson 2017). While considering this point, I shall employ Nussbaum's (1999) argument that merely comparing virtue vs. non-virtue ethics is misleading since Kantian and Utilitarian perspectives also embrace several aspects of virtue ethics. So, anti-Utilitarian (e.g., Kantian) and anti-Kantian (e.g., Utilitarian) perspectives will be examined. Anti-Utilitarians prioritize reason over passion, so they value habits as moral guidance relatively less value. On the other hand, anti-Kantians argue that one's subjective motive and desire dominate reason. These two views, compared with the Neo-Aristotelian perspective, may miss some fundamental aspects of moral functioning. If we solely endorse the Anti-Utilitarian view while designing moral education, habituation- and affect-related aspects, which are necessary for moral development given evidence, may be undervalued. If we take the Anti-Kantian perspective, phronesis-related aspects, which are required for self- and other-understanding and optimal neural learning, could not be well valued. Compared with this view, the Neo-Aristotelian perspective, which underscores the coordination of both aspects for flourishing, is perhaps most consistent with neuroscientific evidence within the context of moral development and education.

So far, I have discussed two central themes proposed in virtue ethics, habituation of virtues and cultivation of phronesis, as qualities to be pursued in moral education with evidence from neuroscience. The evidence may provide useful insights regarding what shall be pursued in moral development and education. However, I am unable to give a complete answer to the question, what shall be the purposes of moral education, at this point. The majority of the introduced neuroimaging studies have addressed general psychological functioning, which is perhaps indirectly associated with morality, instead of moral functioning directly. Particularly, they have not delved into concepts proposed in virtue ethics, (habituation of) virtues and phronesis, so they might be able to provide only partial supports to my points. To understand how neuroscience supports ideas regarding purposes of moral education proposed by virtue ethics better, it would be necessary to conduct additional neuroscientific studies that directly address themes related to virtue ethics, particularly virtue habituation and phronesis.

References

Anscome GEM (1958) Modern Moral Philosophy. Philosophy 33:1–19

Athanassoulis N (2000) A Response to Harman: Virtue Ethics and Character Traits. Proceedings of the Aristotelian Society (Hardback) 100:215–221. https://doi.org/10.1111/j.0066-7372.2003.00012.x

Babyak MA (2004) What You See May Not Be What You Get: A Brief, Nontechnical Introduction to Overfitting in Regression-Type Models. Psychosomatic Medicine 66:411–421. https://doi.org/10.1097/01.psy.0000127692.23278.a9

Bestmann S, Feredoes E (2013) Combined neurostimulation and neuroimaging in cognitive neuroscience: past, present, and future. Annals of the New York Academy of Sciences 1296:11–30. https://doi.org/10.1111/nyas.12110

Blair RJR (2007) The amygdala and ventromedial prefrontal cortex in morality and psychopathy. Trends in cognitive sciences 11:387–92. https://doi.org/10.1016/j.tics.2007.07.003

Bowers JS (2016) The Practical and Principled Problems With Educational Neuroscience. Psychological Review. https://doi.org/10.1037/rev0000025

Bressler SL, Menon V (2010) Large-scale brain networks in cognition: emerging methods and principles. Trends in Cognitive Sciences 14:277–290. https://doi.org/S1364-6613(10)00089-6 [pii] 10.1016/j.tics.2010.04.004

Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. Annals of the New York Academy of Sciences 1124:1–38. https://doi.org/10.1196/annals.1440.011

Bulger SM, Housner LD, Lee AM (2008) Curriculum Alignment: A View from Physical Education Teacher Education. Journal of Physical Education, Recreation & Dance 79:44–49. https://doi.org/10.1080/07303084.2008.10598215

Bzdok D, Schilbach L, Vogeley K, et al (2012) Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Structure and Function 217:783–796. https://doi.org/10.1007/s00429-012-0380-y

Cáceda R, James GA, Ely TD, et al (2011) Mode of Effective Connectivity within a Putative Neural Network Differentiates Moral Cognitions Related to Care and Justice Ethics. PLoS ONE 6:e14730. https://doi.org/10.1371/journal.pone.0014730

Churchland PM (1998) Toward a Cognitive Neurobiology of the Moral Virtues. Topoi 17:83–96

Colby A, Damon W (1993) The uniting of self and morality in the development of extraordinary moral commitment. In: Noam GG, Wren TE (eds) The moral self. MIT Press, Cambridge, MA, pp 149–174

Curren R (1999) Cultivating the Intellectual and Moral Virtues. In: Carr D, Steutel J (eds) Virtue Ethics and Moral Education. Routledge, London, UK, pp 69–83

Cushman F, Kumar V, Railton P (2017) Moral learning: Psychological and philosophical perspectives. Cognition 167:1–10. https://doi.org/10.1016/j.cognition.2017.06.008

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 62:101–129. https://doi.org/10.1159/000496136

De Caro M, Marraffa M, Vaccarezza MS (2021) The priority of phronesis: How to rescue virtue theory from its crisis. In: De Caro M, Vaccarezza MS (eds) Practical Wisdom: Philosophical and Psychological Perspectives. Routledge, New York, NY, pp 29–51

De Caro M, Vaccarezza MS, Niccoli A (2018) Phronesis as Ethical Expertise: Naturalism of Second Nature and the Unity of Virtue. J Value Inquiry 52:287–305. https://doi.org/10.1007/s10790-018-9654-9

Dockès J, Poldrack RA, Primet R, et al (2020) NeuroQuery, comprehensive meta-analysis of human brain mapping. eLife 9:. https://doi.org/10.7554/eLife.53385

Eickhoff SB, Kernbach J, Bzdok D (2020) Meta-Analyses in Basic and Clinical Neuroscience: State of the Art and Perspective. In: Ulmer S, Jansen O (eds) fMRI. Springer International Publishing, Cham, pp 117–129

Epskamp S, Fried EI (2018) A tutorial on regularized partial correlation networks. Psychological Methods 23:. https://doi.org/10.1037/met0000167

Fede SJ, Dean SF, Manuweera T, Momenan R (2020) A Guide to Literature Informed Decisions in the Design of Real Time fMRI Neurofeedback Studies: A Systematic Review. Front Hum Neurosci 14:60. https://doi.org/10.3389/fnhum.2020.00060

Friston KJ, Buechel C, Fink GR, et al (1997) Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6:218–29. https://doi.org/10.1006/nimg.1997.0291

Greene JD (2014) Beyond Point-and-Shoot Morality: Why Cognitive (Neuro)Science Matters for Ethics. Ethics 124:695–726. https://doi.org/10.1086/675875

Greene JD, Sommerville RB, Nystrom LE, et al (2001) An fMRI investigation of emotional engagement in moral judgment. Science 293:2105–2108. https://doi.org/10.1126/science.1062872

Hackel LM, Amodio DM (2018) Computational neuroscience approaches to social cognition. Current Opinion in Psychology 24:92–97. https://doi.org/10.1016/j.copsyc.2018.09.001

Han H (2017) Neural correlates of moral sensitivity and moral judgment associated with brain circuitries of selfhood: A meta-analysis. Journal of Moral Education 46:97–113. https://doi.org/10.1080/03057240.2016.1262834

Han H (2016) How can neuroscience contribute to moral philosophy, psychology and education based on Aristotelian virtue ethics? International Journal of Ethics Education 1:201–217. https://doi.org/10.1007/s40889-016-0016-9

Han H (2022) A Novel Method to Use Coordinate Based Meta-Analysis to Determine a Prior Distribution for Voxelwise Bayesian Second-Level fMRI Analysis. Mathematics 10:356. https://doi.org/10.3390/math10030356

Han H, Chen J, Jeong C, Glover GH (2016) Influence of the cortical midline structures on moral emotion and motivation in moral decision-making. Behavioural Brain Research 302:237–251. https://doi.org/10.1016/j.bbr.2016.01.001

Han H, Kim J, Jeong C, Cohen GL (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, Lee K, Soylu F (2020) Applying the Deep Learning Method for Simulating Outcomes of Educational Interventions. SN Computer Science 1:70. https://doi.org/10.1007/s42979-020-0075-z

Han H, Soylu F, Anchan DM (2019) Connecting Levels of Analysis in Educational Neuroscience: A Review of Multi-level Structure of Educational Neuroscience with Concrete Examples. Trends in Neuroscience and Education 100113. https://doi.org/10.1016/j.tine.2019.100113

Han H, Workman CI, May J, et al (2022) Which moral exemplars inspire prosociality? Philosophical Psychology 1–28. https://doi.org/10.1080/09515089.2022.2035343

Hohenfeld C, Nellessen N, Dogan I, et al (2017) Cognitive Improvement and Brain Changes after Real-Time Functional MRI Neurofeedback Training in Healthy Elderly and Prodromal Alzheimer's Disease. Front Neurol 8:384. https://doi.org/10.3389/fneur.2017.00384

Imperatori C, Della Marca G, Amoroso N, et al (2017) Alpha/Theta Neurofeedback Increases Mentalization and Default Mode Network Connectivity in a Non-Clinical Sample. Brain Topogr 30:822–831. https://doi.org/10.1007/s10548-017-0593-8

Johansen-Berg H (2012) The future of functionally-related structural change assessment. NeuroImage 62:1293–1298. https://doi.org/10.1016/j.neuroimage.2011.10.073

Jordan K, Kristjánsson K (2017) Sustainability, virtue ethics, and the virtue of harmony with nature. Environmental Education Research 23:1205–1229. https://doi.org/10.1080/13504622.2016.1157681

Kim M, Park B, Young L (2020) The Psychology of Motivated versus Rational Impression Updating. Trends in Cognitive Sciences 24:101–111. https://doi.org/10.1016/j.tics.2019.12.001

Klein C (2012) Cognitive Ontology and Region- versus Network-Oriented Analyses. Philos of Sci 79:952–960. https://doi.org/10.1086/667843

Koban L, Gianaros PJ, Kober H, Wager TD (2021) The self in context: brain systems linking mental and physical health. Nat Rev Neurosci 22:309–322. https://doi.org/10.1038/s41583-021-00446-8

Kristjánsson K (2013) Virtues and vices in positive psychology: A philosophical critique. Cambridge University Press, New York, NY

Kristjansson K (2015) Aristotelian Character Education. Routledge, Abington

Kristjánsson K (2005) Smoothing It: Some Aristotelian misgivings about the phronesis-praxis perspective on education. Educational Philosophy and Theory 37:455–473. https://doi.org/10.1111/j.1469-5812.2005.00135.x

Kristjánsson K, Fowers B (2022) Phronesis as moral decathlon: contesting the redundancy thesis about phronesis. Philosophical Psychology 1–20. https://doi.org/10.1080/09515089.2022.2055537

Kristjánsson K, Fowers B, Darnell C, Pollard D (2021) Phronesis (Practical Wisdom) as a Type of Contextual Integrative Thinking. Review of General Psychology 25:239–257. https://doi.org/10.1177/10892680211023063

Lampit A, Hallock H, Suo C, et al (2015) Cognitive training-induced short-term functional and long-term structural plastic change is related to gains in global cognition in healthy older adults: a pilot study. Front Aging Neurosci 7:. https://doi.org/10.3389/fnagi.2015.00014

Lapsley DK (2021) The Developmental Science of Phronesis. In: De Caro M, Vaccarezza MS (eds) Practical Wisdom: Philosophical and Psychological Perspectives. Routledge, New York

Lenzen LM, Donges MR, Eickhoff SB, Poeppl TB (2021) Exploring the neural correlates of (altered) moral cognition in psychopaths. Behavioral Sci & The Law 39:731–740. https://doi.org/10.1002/bsl.2539

Li C-X, Zhang X (2018) Evaluation of prolonged administration of isoflurane on cerebral blood flow and default mode network in macaque monkeys anesthetized with different maintenance doses. Neuroscience Letters 662:402–408. https://doi.org/10.1016/j.neulet.2017.10.034

Lou HC, Luber B, Stanford A, Lisanby SH (2010) Self-specific processing in the default network: a single-pulse TMS study. Experimental brain research 207:27–38. https://doi.org/10.1007/s00221-012-3249-7

McDiarmid AD, Tullett AM, Whitt CM, et al (2021) Psychologists update their beliefs about effect sizes after replication studies. Nat Hum Behav 5:1663–1673. https://doi.org/10.1038/s41562-021-01220-7

Navarini C (2020) The Likelihood of Actions and the Neurobiology of Virtues: Veto and Consent Power. Ethic Theory Moral Prac 23:309–323. https://doi.org/10.1007/s10677-020-10081-4

Noel J (1999) On the Varieties of Phronesis. Educational Philosophy and Theory 31:273–289. https://doi.org/10.1111/j.1469-5812.1999.tb00466.x

Nussbaum MC (1999) Virtue Ethics: A Misleading Category? The Journal of Ethics 3:163–201. https://doi.org/10.1023/A:1009877217694

O'Reilly JX, Jbabdi S, Behrens TEJ (2012) How can a Bayesian approach inform neuroscience?: How can a Bayesian approach inform neuroscience? European Journal of Neuroscience 35:1169–1179. https://doi.org/10.1111/j.1460-9568.2012.08010.x

Poldrack RA, Yarkoni T (2016) From Brain Maps to Cognitive Ontologies: Informatics and the Search for Mental Structure. Annual Review of Psychology 67:587–612. https://doi.org/10.1146/annurev-psych-122414-033729

Prehn K, Korczykowski M, Rao H, et al (2015) Neural Correlates of Post-Conventional Moral Reasoning: A Voxel-Based Morphometry Study. PLOS ONE 10:e0122914. https://doi.org/10.1371/journal.pone.0122914

Pujol J, Batalla I, Contreras-Rodríguez O, et al (2012) Breakdown in the brain network subserving moral judgment in criminal psychopathy. Social Cognitive and Affective Neuroscience 7:917–923. https://doi.org/10.1093/scan/nsr075

Raine A, Yang Y (2006) Neural foundations to moral reasoning and antisocial behavior. Social Cognitive and Affective Neuroscience 1:203–213. https://doi.org/10.1093/scan/nsl033

Ramsey JD, Hanson SJ, Hanson C, et al (2010) Six problems for causal inference from fMRI. NeuroImage 49:1545–1558. https://doi.org/10.1016/j.neuroimage.2009.08.065

Riva P, Manfrinati A, Sacchi S, et al (2019) Selective changes in moral judgment by noninvasive brain stimulation of the medial prefrontal cortex. Cogn Affect Behav Neurosci 19:797–810. https://doi.org/10.3758/s13415-018-00664-1

Rubin TN, Koyejo O, Gorgolewski KJ, et al (2017) Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition. PLoS Comput Biol 13:e1005649. https://doi.org/10.1371/journal.pcbi.1005649

Ruff CC, Ugazio G, Fehr E (2013) Changing Social Norm Compliance with Noninvasive Brain Stimulation. Science 342:482–484. https://doi.org/10.1126/science.1241399

Sämann PG, Wehrle R, Hoehn D, et al (2011) Development of the Brain's Default Mode Network from Wakefulness to Slow Wave Sleep. Cerebral Cortex 21:2082–2093. https://doi.org/10.1093/cercor/bhq295

Sanderse W (2014) An Aristotelian Model of Moral Development. Journal of Philosophy of Education. https://doi.org/10.1111/1467-9752.12109

Sanderse W (2020) Does Aristotle believe that habituation is only for children? Journal of Moral Education 49:98–110. https://doi.org/10.1080/03057240.2018.1497952

Schurz M, Radua J, Tholen MG, et al (2021) Toward a hierarchical model of social cognition: A neuroimaging meta-analysis and integrative review of empathy and theory of mind. Psychological Bulletin 147:293–327. https://doi.org/10.1037/bul0000303

Sinharay S, Stern HS (2002) On the Sensitivity of Bayes Factors to the Prior Distributions. The American Statistician 56:196–201. https://doi.org/10.1198/000313002137

Steutel J, Spiecker B (2004) Cultivating sentimental dispositions through Aristotelian habituation. Journal of Philosophy of Education 38:531–549. https://doi.org/10.1111/j.0309-8249.2004.00403.x

Tymofiyeva O, Gaschler R (2021) Training-Induced Neural Plasticity in Youth: A Systematic Review of Structural and Functional MRI Studies. Front Hum Neurosci 14:497245. https://doi.org/10.3389/fnhum.2020.497245

Valk SL, Bernhardt BC, Trautwein F-M, et al (2017) Structural plasticity of the social brain: Differential change after socio-affective and cognitive mental training. Sci Adv 3:e1700489. https://doi.org/10.1126/sciadv.1700489

Vance J, Werner PJ (2022) Attentional Moral Perception. Journal of Moral Philosophy. https://doi.org/10.1163/17455243-20220001

Yarkoni T, Poldrack RA, Nichols TE, et al (2011) Large-scale automated synthesis of human functional neuroimaging data. Nature Methods 8:665–670. https://doi.org/10.1038/nmeth.1635

Yeager DS, Walton GM (2011) Social-psychological interventions in education: They're not magic. Review of Educational Research 81:267–301. https://doi.org/10.3102/0034654311405999

Young L, Camprodon JA, Hauser M, et al (2010) Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. Proceedings of the National Academy of Sciences of the United States of America 107:6753–6758. https://doi.org/10.1073/pnas.0914826107

Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society Series B: Statistical Methodology 67:301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Supplementary Materials

Methodological Further Details

NeuroSynth and NeuroQuery

NeuroSynth and NeuroQuery, online-based automatized tools to meta-analyze and synthesize large-scale neuroimaging database (Yarkoni et al. 2011; Dockès et al. 2020), allow us to get information about "|Data," brain activity data cross diverse functional domains, and thus, to conduct reverse inference in a more reliable and valid way (Poldrack 2011). Such tools provide researchers with analysis results with an option to explore P(Hypothesis|Data). NeuroSynth generates a reverse inference map demonstrating which regions are particularly associated with a specific keyword entered to the tool while controlling for the baseline brain activity, which is extracted and estimated from large-scale database (Yarkoni et al. 2011). NeuroQuery is an online tool to explore the neural correlates of interest with large-scale online database through machine learning (Dockès et al. 2020). With the machine learning technology, NeuroQuery presents a likelihood map demonstrating which brain regions are most likely to be associated with the entered keyword.

NeuroSynth decoder

For each keyword representing a specific psychological functioning, NeuroSynth already created an activation likelihood image, which demonstrates brain regions showing significant activity associated with the specific psychological functioning after controlling for baseline brain activity. With the aforementioned image database, once a user provides an input image to the decoder, it examines which keywords show the most significant correlation with the input image. Given the decoder provides information about the correlation between psychological functionalities of interest at the neural level after considering the base brain activity across all possible task conditions stored in the large-scale database, it would also be capable of alleviating the reserve inference issue. With the NeuroSynth decoder, I examined which psychological functionalities other than moral functioning showed the most significant correlation with moral functioning at the neural level.

For input images, I created meta-analyzed images reporting the common neural correlates of moral functioning. In this process, I used four different meta-analysis results: the result reported in Han's (2017) meta-analysis[1], NeuroSynth[2] and NeuroQuery[3] with a keyword "moral", and the result of activation likelihood estimation meta-analysis with coordinate information gathered from BrainMap[4] with a keyword "moral." When these four result images were entered to NeuroSynth decoder, interestingly, I was able to see that several keywords reported highest correlation coefficients commonly across four cases. When keywords related to brain regions and circuitries (e.g., "medial prefrontal") were excluded, keywords related to the theory of mind (or mentalizing) and self-referencing (or autobiography) showed highest correlation coefficients in all four cases.

Neurofeedback

One of the most widely tested psychological functionalities with the neurofeedback paradigm was self-regulation. While undergoing the neurofeedback intervention targeting the

---

[1] Decoding result with Han (2017) available at https://neurosynth.org/decode/?neurovault=544822

[2] Decoding result with the NeuroSynth map available at https://neurosynth.org/decode/?neurovault=442007

[3] Decoding result with the NeuroQuery map available at https://neurosynth.org/decode/?neurovault=442779

[4] Decoding result with the BrainMap meta-analysis map available at https://neurosynth.org/decode/?neurovault=505936

improvement of self-regulation, participants were provided with real-time feedback about whether they need to up- or down-regulate their emotional arousal and cognitive response (Koush et al. 2015). The feedback information was formed by activity in brain regions or networks associated with self-regulatory functionalities, such as the DMN regions (McDonald et al. 2017). When activity in such regions or networks of interest deviates from the desirable level, a message requesting participants to up- or down-regulate the functionalities of interest accordingly (Watanabe et al. 2017). Through the neurofeedback training, participants were able to learn how to regulate their emotional and cognitive processes to maintain the appropriate level of arousal.

Results from Previous Neuroscientific Studies

Results from large-scale neuroimaging analyses

When a keyword "moral" was entered to NeuroSynth, the resultant map (https://neurosynth.org/analyses/terms/moral/) demonstrates that brain regions located within the default mode network (DMN), which is associated with social cognition and self-related processes (Han 2017; Koban et al. 2021), were particularly associated with "moral." The similar pattern could also be observed when the keyword "moral" was entered to NeuroQuery (https://neuroquery.org/query?text=moral+).

In addition, I also employed a decoder, an additional feature implemented in NeuroSynth, to explore the neural activity patterns significantly and uniquely associated with moral functioning. NeuroSynth decoder enables us to explore which keywords are most likely to be associated with an entered statistical image (Yarkoni et al. 2011). When keywords related to brain regions and circuitries (e.g., "medial prefrontal") were excluded, keywords related to the theory of mind (or mentalizing) and self-referencing (or autobiography) showed highest correlation coefficients in all cases. The results may suggest that the neural circuitries and psychological processes related to morality and moral functioning are uniquely associated with the theory of mind and self-related processes.

Results from studies employing network-based approaches

There are two widely used analysis methods to examine brain networks of interest, instead of specific brain regions. First, the psychophysiological interaction analysis demonstrates which brain regions are interacting with each other under a specific task condition (Friston et al. 1997). Moreover, connectivity analysis methods have been widely used to examine how activity in different brain regions are associated serially (Cáceda et al. 2011).

In the field of research on neuroscience of morality, there have been several previous neuroimaging studies that employed such network-based approaches. In fact, when the neural correlates of moral functioning were examined in recent studies, their findings suggested that such neural correlates could not be localize in specific brain regions, such as the prefrontal cortex (Moll et al. 2007). Alternatively, they proposed that the correlates were dispersed across the multiple different brain regions. Consistently, functional neuroimaging studies and meta-analyses have also reported that multiple brain regions, particularly those within the DMN, were closely associated with moral functioning (Bzdok et al. 2012; Sevinc and Spreng 2014; Boccia et al. 2016; Garrigan et al. 2016; Han 2017; Eres et al. 2017). Thus, if we solely focus on specific locations of a brain while examining the neural-level nature of morality, then evidence acquired from such attempted could be misleading.

Previous studies employing these methods have demonstrated that brain regions in the DMN are functionally connected with each other in the moral task conditions (Cáceda et al. 2011; Pujol et al. 2012). Furthermore, the DMN regions related to self-related processes were found to be interacting with other brain regions related to cognition, emotion, and motivation within the context of moral functioning (Han et al. 2016).

Results from experimental studies

Several experimental methods in neuroscience at least partially enable us to examine causality between activity in specific brain regions or networks and psychological functionalities of interest. Such methods include the neurostimulation, neurofeedback, and neuroscience-informed intervention (Bestmann and Feredoes 2013). Given education, including moral education, is primarily concerned about how to promote one's development and growth through educational activities (Yeager and Walton 2011), it would be important to address the causality issue if we intend to learn from neuroscience for improving education. Thus, overviewing findings from such experimental studies may provide us with useful insights about mechanisms associated with learning and development, which are fundamental in educational research.

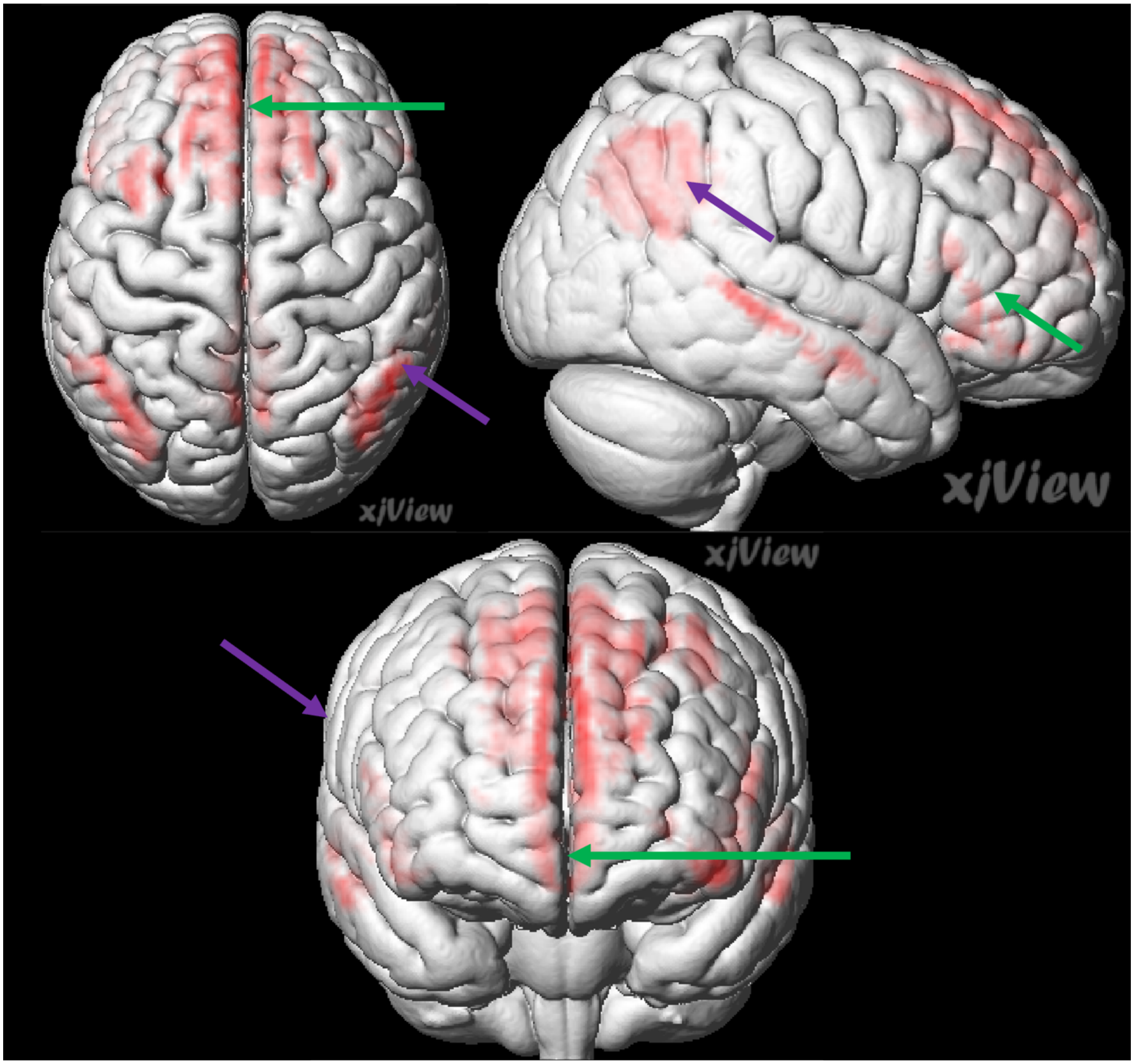
Neurostimulation would be one possible way to examine the causal relation. The Transcranial Magnetic Stimulation (tMS) and Transcranial Direct Current Simulation (tDCS) are the most representative neurostimulation methods (Young et al. 2010; Riva et al. 2019). Because the stimulation methods are non-invasive and non-surgical, they have been widely utilized in neuropsychological studies (Bestmann and Feredoes 2013). These methods allow researchers to alter neural activity in a specific brain region or circuitry with a magnetic field (tMS) or direct electric current (tDCS). Through the stimulation, we can provoke or disrupt brain activity temporally, and we are able to examine a causal relationship between the brain region or network of interest and a psychological functioning of interest (Bestmann and Feredoes 2013).

The neurostimulation methods have been widely employed in neuropsychological studies examining the neural mechanism of moral functioning. For instance, disrupting activity in the prefrontal cortex significantly altered participants' socio-moral decision making and norm compliance (Ruff et al. 2013). Also, when the temporoparietal junction, which is a part of the DMN and associated with mentalizing, was stimulated by the tMS, participants' tendency to infer one's moral intent within moral dilemmas significantly changed (Young et al. 2010). In this study, authors presented different types of stories (intentional harm vs. intentional no harm vs. unintentional harm vs. unintentional no harm) and how participants' judgment (i.e., forbidden vs. permissible) changed when the tMS was applied. In addition, applying the tDCS on the medial prefrontal cortex significantly altered one's deontological versus utilitarian moral judgment (Riva et al. 2019). Riva et al. (2019) employed moral dilemmas in two different types, the Trolley-like and Footbridge-like dilemmas, to examine extent to which participants render utilitarian decisions.

The regions stimulated in these previous studies overlap the regions and circuitries found in the aforementioned large-scale neuroimaging meta-analyses, such as the DMN (Lou et al. 2010) (see Figure S1 for information about whether the tMS and tDCS were applied in the two studies). The evidence from neurostimulation studies may support the causal relations between activity in the stimulated areas and moral functioning. Thus, we may assume that psychological and behavioral processes within the moral domain might be moderated and regulated by neural activity in the aforementioned regions and circuitries.

Figure S1

Default mode network and regions targeted in Riva et al. (2019) and Young et al. (2010).



Note. Red: the default mode network identified in Doucet et al. (2019). Purple: the medial prefrontal cortex region stimulated by the tDCS in Riva et al. (2019). Green: the temporoparietal junction stimulated by the tMS in Young et al. (2010). 3D images generated with xjview (Cui et al. 2015).

One point that we may consider is that these previous studies targeted brain regions associated with moral functioning, particularly Riva et al. (2019) and Young et al. (2010), actually altered such a functioning, not general cognitive functioning. That says, the significantly altered responses from participants during the moral task conditions were not necessarily attributable to disruption of general cognitive capacities, which may be independent from moral functioning. For instance, in Riva et al. (2019), the examine the association between extent to which participants' utilitarian decision-making tendency was altered by the tDCS and their demonological intent. If the disruption of the region of interest influenced general cognition rather than cognition specific within the moral domain, then, the aforementioned tendency should be consistent regardless of one's own deontological intent. Furthermore, Young et al. (2010) demonstrated that participants' mentalizing was differently altered by the tMS across different task conditions. The greater change occurred when presented stories implied intentional harm compared with when stories were about either intentional no-harm, unintentional harm, or unintentional no-harm. Given intent and harm are both fundamental aspects within the moral domain (Cushman and Young 2009), the result might support the same point that the brain stimulation resulted in the disruption of moral functioning instead of domain-general cognitive functioning. Similar to the case of Riva et al. (2019), if the tMS altered general cognitive capacities, then the significant rate changes should also be observed in other conditions as well.

Neurofeedback studies can also provide additional evidence supporting causality. Neurofeedback applies real-time neuroimaging techniques to provide participants with information about how to regulate targeted psychological processes (Fede et al. 2020; see supplmentary materials for methodological further details). Previous studies employing neurofeedback reported not only significant changes in activity in brain regions associated with aimed functionalities during the intervention session, but also long-term improvement in such functionalities (Hohenfeld et al. 2017). The findings may suggest that if we can properly identify brain regions or circuitries associated with psychological functionalities of interest, we can improve the functionalities even in a long-term through neurofeedback. Thus, it would be a possible way to support a causal relation between brain regions or circuitries and psychological functionalities.

Furthermore, evidence from neuroscience-informed intervention studies aiming at promoting specific behavioral outcomes, such as moral behavior, can also be considered. For instance, Han et al.'s (2017) moral exemplar intervention study was informed by meta-analysis of fMRI studies focusing on moral functioning. Based on the meta-analysis finding that self-related brain regions, the DMN in particular (Han 2017), are significantly associated with diverse moral functioning, Han et al. (2017) hypothesized that the perceived connectivity between participants and presented exemplars would affect motivational outcomes. They found that close moral exemplars were more effective in motivating students compared with distant exemplars. The finding may support a point that an understanding of brain functioning can be used to inform educational interventions aiming at promoting desired behavioral outcomes. If such neuroscience-informed interventions work as intended, the fact may at least partially support a causal relation between brain activity and psychological functioning. Plus, it may also suggest that neuroscience can inform education, including moral education.

Additionally, trainings and interventions aiming at behavioral changes have also been known to induce changes in the structure and functioning at the neural level (Johansen-Berg 2012). It may support the point that brain activity and psychological or behavioral functioning may causally and reciprocally influence each other. For instance, a longitudinal neuroimaging study demonstrated that participation in long-term training programs addressing various aspects of socio-affective and cognitive processes (e.g., mindfulness, compassion, perspective taking) induced significant structural changes in brain regions associated with the targeted processes (Valk et al. 2017). Across diverse domains of affective, cognitive, and motivational processes, psychological training programs have significantly affected the brain structure, connectivity, and functioning in regions associated with the aimed processes (Lampit et al. 2015; Tymofiyeva and Gaschler 2021). Such evidence unequivocally suggests that well-designed trainings are capable of inducing changes at the motivational, behavioral, and even neural levels when the trainings are performed in a long term.

Bayes rule

That says, we, humans, have prior beliefs (H) in a specific domain, which is being updated by observed data (D). Once data is observed, the prior beliefs are updated into posterior beliefs following Bayes Theorem (Han and Park 2018):

$$P(H|D) = \frac{P(H)P(D|H)}{P(D)}$$

where P(H) is the probability distribution of prior beliefs and P(H|D) is that of posterior beliefs. When we observe data, D, which is consistent with our prior beliefs, H, then the posterior beliefs are intensified accordingly. On the other hand, if D contradicts H, then the posterior beliefs become weaker.

Supplementary References

Bestmann S, Feredoes E (2013) Combined neurostimulation and neuroimaging in cognitive neuroscience: past, present, and future. Ann N Y Acad Sci 1296:11–30. https://doi.org/10.1111/nyas.12110

Boccia M, Dacquino C, Piccardi L, et al (2016) Neural foundation of human moral reasoning: an ALE meta-analysis about the role of personal perspective. Brain Imaging Behav. https://doi.org/10.1007/s11682-016-9505-x

Bzdok D, Schilbach L, Vogeley K, et al (2012) Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct 217:783–796. https://doi.org/10.1007/s00429-012-0380-y

Cáceda R, James GA, Ely TD, et al (2011) Mode of Effective Connectivity within a Putative Neural Network Differentiates Moral Cognitions Related to Care and Justice Ethics. PLoS ONE 6:e14730. https://doi.org/10.1371/journal.pone.0014730

Cui X, Li J, Song X (2015) xjview. http://www.alivelearn.net/xjview. Accessed 28 Jun 2015

Cushman F, Young L (2009) The Psychology of Dilemmas and the Philosophy of Morality. Ethical Theory Moral Pract 12:9–24. https://doi.org/10.1007/s10677-008-9145-3

Dockès J, Poldrack RA, Primet R, et al (2020) NeuroQuery, comprehensive meta-analysis of human brain mapping. eLife 9:. https://doi.org/10.7554/eLife.53385

Doucet GE, Lee WH, Frangou S (2019) Evaluation of the spatial variability in the major resting-state networks across human brain functional atlases. Hum Brain Mapp 40:4577–4587. https://doi.org/10.1002/hbm.24722

Eres R, Louis WR, Molenberghs P (2017) Common and distinct neural networks involved in fMRI studies investigating morality: an ALE meta-analysis. Soc Neurosci 1–15. https://doi.org/10.1080/17470919.2017.1357657

Fede SJ, Dean SF, Manuweera T, Momenan R (2020) A Guide to Literature Informed Decisions in the Design of Real Time fMRI Neurofeedback Studies: A Systematic Review. Front Hum Neurosci 14:60. https://doi.org/10.3389/fnhum.2020.00060

Friston KJ, Buechel C, Fink GR, et al (1997) Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6:218–29. https://doi.org/10.1006/nimg.1997.0291

Garrigan B, Adlam ALR, Langdon PE (2016) The neural correlates of moral decision-making: A systematic review and meta-analysis of moral evaluations and response decision judgements. Brain Cogn 108:88–97. https://doi.org/10.1016/j.bandc.2016.07.007

Han H (2017) Neural correlates of moral sensitivity and moral judgment associated with brain circuitries of selfhood: A meta-analysis. J Moral Educ 46:97–113. https://doi.org/10.1080/03057240.2016.1262834

Han H, Chen J, Jeong C, Glover GH (2016) Influence of the cortical midline structures on moral emotion and motivation in moral decision-making. Behav Brain Res 302:237–251. https://doi.org/10.1016/j.bbr.2016.01.001

Han H, Kim J, Jeong C, Cohen GL (2017) Attainable and Relevant Moral Exemplars Are More Effective than Extraordinary Exemplars in Promoting Voluntary Service Engagement. Front Psychol 8:283. https://doi.org/10.3389/fpsyg.2017.00283

Han H, Park J (2018) Using SPM 12's second-level bayesian inference procedure for fMRI analysis: Practical guidelines for end users. Front Neuroinformatics 12:1. https://doi.org/10.3389/fninf.2018.00001

Hohenfeld C, Nellessen N, Dogan I, et al (2017) Cognitive Improvement and Brain Changes after Real-Time Functional MRI Neurofeedback Training in Healthy Elderly and Prodromal Alzheimer's Disease. Front Neurol 8:384. https://doi.org/10.3389/fneur.2017.00384

Johansen-Berg H (2012) The future of functionally-related structural change assessment. NeuroImage 62:1293–1298. https://doi.org/10.1016/j.neuroimage.2011.10.073

Koban L, Gianaros PJ, Kober H, Wager TD (2021) The self in context: brain systems linking mental and physical health. Nat Rev Neurosci 22:309–322. https://doi.org/10.1038/s41583-021-00446-8

Koush Y, Meskaldji D-E, Pichon S, et al (2015) Learning Control Over Emotion Networks Through Connectivity-Based Neurofeedback. Cereb Cortex bhv311. https://doi.org/10.1093/cercor/bhv311

Lampit A, Hallock H, Suo C, et al (2015) Cognitive training-induced short-term functional and long-term structural plastic change is related to gains in global cognition in healthy older adults: a pilot study. Front Aging Neurosci 7:. https://doi.org/10.3389/fnagi.2015.00014

Lou HC, Luber B, Stanford A, Lisanby SH (2010) Self-specific processing in the default network: a single-pulse TMS study. Exp Brain Res 207:27–38. https://doi.org/10.1007/s00221-012-3249-7

McDonald AR, Muraskin J, Dam NTV, et al (2017) The real-time fMRI neurofeedback based stratification of Default Network Regulation Neuroimaging data repository. NeuroImage 146:157–170. https://doi.org/10.1016/j.neuroimage.2016.10.048

Moll J, De Oliveira-Souza R, Garrido GJ, et al (2007) The self as a moral agent: linking the neural bases of social agency and moral sensitivity. Soc Neurosci 2:336–352. https://doi.org/10.1080/17470910701392024

Poldrack RA (2011) Inferring mental states from neuroimaging data: From reverse inference to large-scale decoding. Neuron 72:692–697. https://doi.org/10.1016/j.neuron.2011.11.001

Pujol J, Batalla I, Contreras-Rodríguez O, et al (2012) Breakdown in the brain network subserving moral judgment in criminal psychopathy. Soc Cogn Affect Neurosci 7:917–923. https://doi.org/10.1093/scan/nsr075

Riva P, Manfrinati A, Sacchi S, et al (2019) Selective changes in moral judgment by noninvasive brain stimulation of the medial prefrontal cortex. Cogn Affect Behav Neurosci 19:797–810. https://doi.org/10.3758/s13415-018-00664-1

Ruff CC, Ugazio G, Fehr E (2013) Changing Social Norm Compliance with Noninvasive Brain Stimulation. Science 342:482–484. https://doi.org/10.1126/science.1241399

Sevinc G, Spreng RN (2014) Contextual and perceptual brain processes underlying moral cognition: A quantitative meta-analysis of moral reasoning and moral emotions. PLoS ONE 9:e87427. https://doi.org/10.1371/journal.pone.0087427

Tymofiyeva O, Gaschler R (2021) Training-Induced Neural Plasticity in Youth: A Systematic Review of Structural and Functional MRI Studies. Front Hum Neurosci 14:497245. https://doi.org/10.3389/fnhum.2020.497245

Valk SL, Bernhardt BC, Trautwein F-M, et al (2017) Structural plasticity of the social brain: Differential change after socio-affective and cognitive mental training. Sci Adv 3:e1700489. https://doi.org/10.1126/sciadv.1700489

Watanabe T, Sasaki Y, Shibata K, Kawato M (2017) Advances in fMRI Real-Time Neurofeedback. Trends Cogn Sci 21:997–1010. https://doi.org/10.1016/j.tics.2017.09.010

Yarkoni T, Poldrack RA, Nichols TE, et al (2011) Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670. https://doi.org/10.1038/nmeth.1635

Yeager DS, Walton GM (2011) Social-psychological interventions in education: They're not magic. Rev Educ Res 81:267–301. https://doi.org/10.3102/0034654311405999

Young L, Camprodon JA, Hauser M, et al (2010) Disruption of the right temporoparietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. Proc Natl Acad Sci U S A 107:6753–6758. https://doi.org/10.1073/pnas.0914826107

---

[i] Kristjánsson and Fowers (2022) employed the example of decathlon to explain how the concept of *phronesis*, practical wisdom, plays fundamental roles in optimal moral functioning. Given *phronesis* is associated with how to render the best decision in a given situation, it is inseparable from considering multiple complex conditions as well as organizing the coordination of virtues and values. These aspects of phronesis are similar to the nature of decathlon, which requires its players to well organize cooperation of diverse athletic skills across different subjects depending on situational factors instead of merely focusing on individual subjects.