The Stability and Variability of Values: A Cognitive Neuroscience Perspective

Authors: Yue Tong, Huang Xiting, Xu Ying, Pan Sicun, Huang Xiting

Date: 2020-08-28T00:00:00+00:00

Abstract

The stability and variability of values has long been a focal point of debate in this field. In recent years, researchers have begun to explore the mechanisms underlying differences in value stability from a cognitive neuroscience perspective. Existing neural response evidence suggests that when a value is cognitively constructed based on deontological absolute rules, or when it exhibits representational consistency with self-concept, it tends to remain stable; otherwise, it may undergo changes in importance under external influences (such as persuasion from others). Future research should further enrich and refine cognitive neuroscientific processing models of the relationship between value stability and variability, and explore the cognitive neural mechanisms underlying long-term value change, to facilitate the practical application of theoretical research in values education.

Full Text

Preamble

The Stability and Variability of Values: A Cognitive Neuroscience Perspective

YUE Tong, HUANG Xiting, XU Ying, PAN Sicun
(Research Center for Psychology and Social Development, Southwest University; School of Psychology, Southwest University, Chongqing 400715, China)

Abstract

The stability and variability of values have long been central to debate in this field. In recent years, researchers have begun exploring the mechanisms underlying differences in value stability from a cognitive neuroscience perspective. Existing neurophysiological evidence indicates that when a value is cognitively

constructed based on deontological absolute rules or exhibits representational consistency with self-concept, it tends to remain stable. Otherwise, it may undergo significant changes in importance under external influence (e.g., persuasion). Future research should further enrich and refine cognitive neuroscientific processing models that elucidate the relationship between value stability and variability, and explore the cognitive neural mechanisms behind long-term value change to promote the practical application of theoretical research in values education.

Keywords: values; stability; variability; cognitive neural mechanisms

1. Introduction

Values represent a belief system through which people judge right from wrong, good from evil, and gain from loss. They not only guide us in pursuing our ideals but also determine the choices we make in life (Huang, 2014). Psychology has long engaged with the study of values, conducting extensive research on their psychological structure, processing mechanisms, functional roles, and measurement tools, yielding substantial progress in related domains (Huang & Zheng, 2005). Values are generally considered relatively stable psychological structures (Rokeach, 1973), with many scholars arguing that once formed, values can only undergo minor adjustments within existing value structures (Lubinski, Schmidt, & Benbow, 1996; Schwartz & Rubel, 2005). However, both everyday experience and existing research findings (Rokeach, 1973; Sheldon, 2005) demonstrate that certain values can change substantively. The question of whether values are stable or variable has remained a focal point of contention among researchers.

Research on value stability has primarily focused on two aspects. First, within Schwartz' s theoretical framework of values, longitudinal studies using instruments such as the Schwartz Value Survey (SVS; Schwartz, 1992) and the Portrait Values Questionnaire (PVQ; Schwartz, 2001) have revealed trends in value stability across time and contexts, finding that individuals' value rankings remain largely stable (Milfont et al., 2016; Sundberg, 2016). Second, researchers have examined the most stable components of values, such as protected values (value concepts that resist utilitarian trade-offs) (Gibson et al., 2016), investigating how they withstand challenges from temptation or social influence (Atran & Ginges, 2015). Overall, extensive research indicates that people' s value hierarchies remain generally stable, with certain values effectively resisting external informational impact and exhibiting absolutely stable psychological characteristics.

Nevertheless, numerous studies have also found that certain values can change under specific situational or experimental manipulations. Bardi and Goodwin (2011) defined value change as alterations in the importance of particular values, identifying five methods that can induce such change: priming (e.g., bicultural individuals exhibit corresponding values based on cultural cues; Gardner et al., 1999), adaptation (e.g., immigrants'values gradually converge with those of their

host country while diverging from their birthplace values; Lönnqvist et al., 2013; Rudnev, 2014), identification (e.g., parenthood triggers value shifts, with new parents emphasizing conservation values more strongly; Lönnqvist et al., 2018), consistency maintenance (e.g., changing behavior first, then adjusting values to rationalize the behavior; Grant & Sagiv, 2009), and persuasion (e.g., providing adolescents with persuasive facts increases the importance of stimulation, self-direction, and universalism values while decreasing security and conformity values; Döring & Hillbrink, 2015).

In summary, previous evidence demonstrates that human values possess both relatively stable psychological structures and situational malleability. This suggests that discussing values as either stable or variable is overly simplistic. Instead, values should be viewed as an open, dynamic system (Huang, 2014) in which some values tend to maintain stability while others are more susceptible to external influence during continuous interaction with environmental information. However, prior research has largely described the phenomena of value stability and change without exploring underlying mechanisms or attempting to resolve the stability-variability debate.

Recent advances in cognitive neuroscience and functional magnetic resonance imaging (fMRI) have made it possible to reveal the psychological processing and neural mechanisms of values, providing more direct analytical perspectives for understanding value stability and variability. This paper reviews relevant research in cognitive neuroscience, addressing two central questions: how certain values maintain stability and how others undergo change. We attempt to integrate the stability-variability debate at the level of neural processing mechanisms and conclude by proposing future directions to provide theoretical reference for subsequent research.

2. How Values Maintain Stability

Protected values are considered the most stable components of the value system. Examining their neural processing patterns can illuminate why certain values remain stable. This section uses protected values as a representative case to discuss the cognitive-neural mechanisms underlying value stability from two perspectives: cognitive construction principles and relationship with self-concept.

2.1 Deontological-Based Cognitive Construction

Values are essentially abstract conceptual representations (Huang, 2014) and collections of social rules (Zahn et al., 2007). People process values through two approaches: deontological or utilitarian principles. The former emphasizes judgment of an act's nature rather than its outcomes (Baron, 2009), involving binary right/wrong distinctions similar to moral intuition. The latter operates within a utility framework, weighing pros and cons to maximize benefits. Research indicates that protected values employ deontological rather than utilitarian con-

struction. For instance, Berns et al. (2012) found that protected values activated the temporoparietal junction (TPJ) and ventrolateral prefrontal cortex (vlPFC) —regions associated with moral judgment of beliefs (Young et al., 2010) and semantic rule extraction (Sharot et al., 2012). In contrast, non-protected values more strongly activated the inferior parietal lobules (IPL), which participate in utilitarian value computation and decision-making (Bhanji et al., 2010; Kable & Glimcher, 2009). This suggests that protected values are processed as absolute rules rather than through cost-benefit analysis.

Building on this, Hamid et al. (2019) examined how protected values occupy central positions in belief systems among supporters of Lashkar-e-Toiba (a South Asian terrorist organization). They found that when reporting willingness to sacrifice for protected values, participants showed significantly reduced dorsolateral prefrontal cortex (dlPFC) activity. Since dlPFC is involved in cognitive control and response inhibition (Carter & Veen, 2007), this suggests that protected values involve less deliberation and more automatic deontological processing based on right/wrong distinctions.

Deontological rule construction enables protected values to maintain stability against social conformity pressure or temptation. In Pincus et al.'s (2014) study, participants first rated their agreement with value statements, then learned how many others disagreed with them. When participants maintained their protected values under peer pressure, vlPFC activity increased significantly. Since vlPFC is associated with semantic rule extraction (Sharot et al., 2012), this suggests individuals rely on deontological principles to resist social influence—treating protected values as absolutely correct and non-negotiable. Similarly, Dogan et al. (2016) investigated how individuals maintain integrity values when facing monetary temptation. Participants role-played as CEOs choosing between honest reporting and fraudulent reporting for personal gain. Results showed that individuals who valued integrity more strongly exhibited greater activation in dlPFC and inferior frontal gyrus (IFG) when facing economic temptation. Since IFG is involved in semantic rule representation (Souza et al., 2009) and dlPFC in cognitive control and response inhibition (Carter & Veen, 2007), these findings suggest that individuals who strongly value integrity may extract absolute rule representations and engage cognitive control to resist monetary temptation, thereby maintaining their principles.

In summary, the stability of protected values is closely related to their deontological cognitive construction, which makes them highly principled and effectively resistant to environmental impact.

2.2 Representational Consistency with Self-Concept

Recent neuroimaging research on protected values has also revealed their high representational consistency with self-concept. In Brosch et al.' s (2012) study, participants completed two tasks: rating the personal importance of values and rating their liking of behaviors. Compared to economic value judgments, value

judgments activated the medial prefrontal cortex (mPFC)—a region closely associated with self-related information processing. Additionally, value importance ratings correlated with precuneus activation, which participates in self-related episodic memory and self-reflection (Fransson & Marrelec, 2008). These results preliminarily suggest that important values overlap representationally with self-concept.

Further evidence comes from Kaplan et al. (2016), who used ecologically valid materials (online stories representing different value orientations) to examine neural representation differences between protected and general values. They found that understanding protected values activated posterior medial cortices (PMC, including precuneus and posterior cingulate), TPJ, and mPFC more strongly than non-protected values. These regions have been implicated in autobiographical memory and self-information processing (Araujo et al., 2014), further indicating that protected value processing requires greater involvement of the self-processing system and exhibits high representational consistency with self-concept.

From a functional perspective, the close connection between protected values and self-representation may serve as an important stability-maintaining strategy. The self occupies the core of the personality system, integrating cognition, emotion, motivation, and volitional behavior (Guo, 2018). People show positive affective preferences for self-related information (Falk & Scholz, 2018), making values integrated into self-concept more likely to receive attention. Verplanken and Holland (2002) found that values' influence on behavior depends on their degree of self-internalization—only self-identified values are important and directly influence behavior. Subsequent behavioral research confirmed that priming and highlighting the self significantly strengthens values' predictive power for corresponding behavior (Ehret, 2019; Tao & Au, 2014). Pretus et al.' s (2019) neuroimaging study also showed that stronger willingness to fight or die for a value correlated with greater ventromedial prefrontal cortex (vmPFC) activity. The vmPFC is involved in both subjective value assignment during decision-making (Bartra et al., 2013) and self-related information processing (Yang, Huang, & Fu, 2012), while dlPFC is associated with negotiation reasoning and cost-benefit integration. These findings suggest that self-internalization may lead individuals to assign highest subjective value to a value concept, reducing cognitive trade-offs during processing and making it difficult to challenge. Conversely, when a value becomes part of self-concept and identity (Hitlin, 2003), changing it would mean altering self-schemas—a psychologically costly process that may trigger cognitive dissonance (Grant & Sagiv, 2009), potentially explaining why protected values resist change.

2.3 Summary

Analysis of neuroimaging research on protected values reveals psychological mechanisms underlying value stability. First, protected values are constructed through deontological principles, making them unchangeable "absolute truths"

that bypass cost-benefit analysis and logical reasoning when challenged, thus ensuring stability. Second, protected values exhibit substantial representational overlap with the self-concept system, making values integrated into the self-system more strongly motivated to maintain self-consistency and subjectively assigned highest value, further enhancing stability. These mechanisms provide cognitive-neural perspectives and evidence for understanding value stability while suggesting potential pathways for value change: if values are not constructed as "absolute" rules or integrated into self-concept, might they be more susceptible to change? The following section addresses this question.

3. How Values Change

Although many methods can effectively alter individuals' values, cognitive neuroscience research has primarily focused on persuasion—changing values by increasing the importance of new value concepts or reducing commitment to existing ones. This section uses persuasion as a representative case to discuss the underlying cognitive processing patterns.

3.1 Value Trade-offs and Cognitive Reappraisal

As previously discussed, protected values maintain stability partly through deontological construction as absolute rules, making them resistant to change. However, not all values are constructed deontologically. Utilitarian principles represent an alternative construction method, where individuals calculate expected benefits of different value orientations and choose the most adaptive option when values conflict (Brosch & Sander, 2013). Persuasion works by demonstrating the advantages of new concepts or shortcomings of existing ones, enabling individuals to conduct value trade-offs and make optimal decisions (Falk & Scholz, 2018). Neuroimaging studies show that accepting new values during persuasion activates brain regions responsible for value evaluation and computation. For example, Cooper et al. (2017) found that among sedentary individuals, functional connectivity strength between value computation systems predicted subsequent behavior change: stronger activation and connectivity between vmPFC and ventral striatum (VS) in response to "increase physical activity" messages correlated with greater reductions in sedentary time. Since vmPFC and VS play important roles in utilitarian computation for maximizing primary (e.g., food, sex) and secondary (e.g., money) values (Bartra et al., 2013; McNamee et al., 2013), Cooper et al. (2017) suggested that health concepts may lead some sedentary individuals to recognize health benefits, weigh them in the brain, and alter their lifestyle principles based on self-interest maximization. Similarly, research on smokers found that stronger vmPFC-VS connectivity after exposure to negative smoking information predicted greater likelihood of behavior change (Falk & Scholz, 2018).

Additionally, values contain unreflective cognitive components that people struggle to justify, treating them simply as social rules to follow (Chen & Zheng,

2010). When values are not constructed as deontological "absolute truths," logical persuasion may enable individuals to recognize their unreasonable aspects and alter their value structures through cognitive reappraisal (Bardi & Goodwin, 2011). Yomogida et al. (2017) examined this neural process by having participants rate their agreement with values (e.g., "talking on public transportation") and knowledge beliefs (e.g., "sugar is harmful to health"), then performing four persuasion tasks (agree/disagree with values/knowledge). After persuasion, participants re-rated their agreement. Results showed that value change activated temporal pole and dorsomedial prefrontal cortex more than knowledge belief change—regions important for social information processing and mental state inference (Mars et al., 2012). Additionally, increased opposition to values correlated with enhanced left middle temporal gyrus (MTG) activity, suggesting reasoning and cognitive reappraisal processes (Buhle et al., 2014; Wei et al., 2012). These findings indicate that after receiving persuasive information, individuals engage in judgment, reasoning, and decision-making processes that may deem certain values unreasonable, leading to cognitive deconstruction and reduced endorsement. The capacity to reflect on and reappraise values within a rational framework is essential for value change, suggesting that changeable values are "relative truths" rather than deontological absolute rules.

3.2 Enhancing Association with Self-Concept

As previously noted, integrating values into self-concept increases their importance and endorsement, representing a key mechanism for protected value stability. However, not all values undergo this self-identification process, and values excluded from self-concept are relatively less important. Can persuasion enhance the association between values outside self-identification and self-concept, thereby increasing their importance and changing behavior?

Because direct persuasion to accept previously unendorsed values can trigger self-threat, researchers often combine it with self-affirmation—affirming the self' s wholeness and worth to buffer this threat (Sherman & Hartson, 2011; Wakslak & Trope, 2009; He & Huang, 2012)—to facilitate smoother integration between values and self-concept. This approach has been applied in education and health domains to change values and behaviors (Epton et al., 2015; Sweeney & Moyer, 2015). Recently, Falk et al. (2015) examined this process neurally by recruiting sedentary participants (who did not value healthy lifestyles) and dividing them into experimental and control groups. The experimental group completed a self-affirmation exercise (reflecting on their most important values). Both groups received persuasive messages about health importance and sedentary threats while neural activity was recorded. Results showed that compared to controls, the experimental group significantly reduced sedentary behavior, with reduction magnitude correlating with vmPFC activity. Since vmPFC is a convergence zone for value computation and self-related thinking (Bartra et al., 2013; Levy & Glimcher, 2012), this activation suggests a "self-value integration" process where persuasive concepts are incorporated into self-concept

(Vezich et al., 2017). Falk et al. (2015) concluded that self-affirmation makes health values more self-relevant, facilitating their integration into self-concept and promoting acceptance of corresponding lifestyles. This demonstrates that persuading individuals to incorporate previously unimportant values into their self-system significantly enhances value importance and endorsement, representing an effective method for changing value structures.

3.3 Summary

Persuasion is the most common laboratory method for changing values. Summarizing relevant cognitive neuroscience research reveals its underlying mechanisms: First, when persuaded values are not deontologically constructed (e.g., health concepts are not principled values for most people), persuasion enables utilitarian thinking and optimal choice, increasing value importance. Second, persuasion can trigger cognitive deconstruction, reducing the perceived reasonableness of existing values and weakening commitment. Third, persuasion can incorporate values into self-concept through self-affirmation, increasing their importance and behavioral impact.

4. Understanding the Relationship Between Value Stability and Variability from a Cognitive Neuroscience Perspective

In conclusion, value stability and variability are not mutually exclusive but should be discussed within an integrated system explaining why some values tend toward stability while others are more changeable. Many researchers view values as having hierarchical structures (Brosch et al., 2011; Ehret, 2019; Schwartz, 2012), where values differ in importance: more central values have higher importance and endorsement, and different importance levels involve distinct cognitive processing patterns. Within this framework and incorporating recent cognitive neuroscience evidence, we propose a preliminary cognitive neuroscientific processing model of value stability and variability (see Figure 1 [Figure 1: see original paper]).

First, the value system has hierarchical properties, with different levels showing distinct neural representations. The most important values, such as protected values, are represented as deontological absolute rules involving TPJ, vlPFC, and IFG. Additionally, protected values are neurally linked to the self-processing system, involving mPFC, precuneus, and vmPFC. As value importance decreases, general values may become collections of social concepts and rules requiring recruitment of brain regions associated with abstract conceptual knowledge representation, such as the anterior temporal lobe (aTL) (Duc et al., 2013; Zahn et al., 2009). Second, these neural structural differences produce stability variations in resisting external impact. Protected values' deontological representation confers absolute correctness, making them difficult to replace. Self-internalization makes them supreme values subjectively assigned extremely high worth, rendering external value evaluations difficult to challenge. Con-

versely, general values' neural structures make them unstable. Experimental interventions reveal three pathways for change: (1) cognitive deconstruction and rule reorganization through persuasion, involving TPJ, dmPFC, and MTG in judgment, reasoning, and decision-making, reducing existing values' importance; (2) utilitarian trade-offs between old and new values (activating IPL, dlPFC) leading to acceptance of persuasive information based on self-interest maximism (activating vmPFC, VS), abandoning original values; and (3) self-affirmation activating vmPFC to produce self-value integration, incorporating previously non-self values into self-concept, increasing their importance and stability. In summary, values constructed more through absolute rules and self-internalization show higher stability; deviation from these neural construction patterns reduces importance and increases variability.

[Figure 1: see original paper]

However, this model remains preliminary, based on limited existing evidence with many assumptions requiring further verification. Future research should integrate more value types and influencing factors to enrich and refine the model and enhance its explanatory power.

5. Summary and Outlook

Values are core constructs in individual socialization (Li & Huang, 2013). Stable values provide sustained motivation for goal achievement, while changing inappropriate concepts and behaviors represents crucial work for education and social media. Therefore, revealing the relationship between value stability and variability and exploring underlying mechanisms holds significant theoretical and practical value. Although burgeoning brain mechanism research provides new perspectives and evidence, this field remains in its infancy with notable limitations. Future research should address the following issues.

First, the formation mechanisms of hierarchical value representations need clarification. Current findings suggest the value system has hierarchical properties with differential stability and variability. How does this hierarchy form? No clear answers exist yet, though the process may trace to value formation (Boer & Boehnke, 2016). For example, values vertically transmitted from parents during childhood or strongly endorsed religious values promoted by religious groups often contain absolute "must" components closely linked to self-identity. These deeply rooted behavioral guidelines may constitute initial components of core-level values, differing in stability from general values formed during socialization, but whether different formation pathways exist requires further investigation. Additionally, while absolutely stable values differ from changeable values in rule construction and self-internalization, how these differences are distributed and interact in neural networks, whether correspondences exist between them, and how such correspondences contribute to hierarchical value formation remain unanswered questions requiring more sophisticated designs and techniques.

Second, research scope on value stability and variability should expand. Current stability research focuses almost exclusively on protected values, while change research primarily uses persuasion. This limited scope only roughly outlines cognitive processing mechanisms, obscuring the overall pattern. Many laboratory methods beyond persuasion can change values (Arieli et al., 2014; Bardi & Goodwin, 2011; Hitlin & Civettini, 2017), likely involving distinct psychological pathways and neural mechanisms. Furthermore, most research addresses "cold" cognitive aspects of value stability and change, while "hot" emotional processing may also be involved. For example, when monetary compensation forces compromise on important values, individuals experience negative emotions (e.g., anger, disgust) that strengthen original value commitment—the "backfire effect" (Atran & Ginges, 2012; Sheikh et al., 2012). Although Duc et al. (2013) identified neural correlates of the backfire effect (amygdala, anterior temporal lobe, fronto-temporal-limbic network), how "cold" and "hot" processes interact remains unclear. Future research should address these issues to enrich theoretical models of value hierarchy and cognitive-neural mechanisms, providing more comprehensive understanding of value stability and variability.

Third, cognitive neural mechanisms of long-term value change require exploration. Most current research examines short-term changes in laboratory settings, investigating state-level neural patterns. It remains uncertain whether experimental interventions produce genuine value change, making identified neural activation patterns potentially incomplete. Longitudinal studies show that major life experiences such as migration (Bardi et al., 2014; Lönnqvist et al., 2011) or workplace transitions (Sandal & Bye, 2015; Sandal et al., 2011) induce long-term value changes. Existing values have stable neural foundations reflected in spontaneous brain activity and even structural features (Zacharopoulos et al., 2017; Zacharopoulos et al., 2016). Given brain plasticity, structural changes in value systems must correspond to significant shifts in neural response patterns. Future research should adopt static perspectives, comparing spontaneous neural activity and structural changes before and after value change, to provide stronger evidence for the neural mechanisms of value change.

Finally, intervention studies based on enriched theory can promote practical application in values education. Enhancing the importance and stability of core values and transforming them into emotional identification and behavioral habits represents a crucial task in education and propaganda (Jiang, 2018). However, Yu and Wang (2013) found that although Chinese college students endorse socialist core values, they fail to internalize them. As discussed, self-affirmation methods that integrate values with self-concept have shown significant effects in laboratory studies and are being applied in values and behavior change research abroad. Future values education practice should incorporate relevant theoretical findings with further innovation and expansion to promote and consolidate the stability of socialist core values in individuals' belief systems.

References

Chen, Y., & Zheng, Y. (2010). The consistency controversy between values and behavior. *Advances in Psychological Science*, 18(10), 1612-1619.

He, Y., & Huang, X. (2012). The underlying mechanisms and influencing factors of self-affirmation. *Journal of Southwest University (Social Sciences Edition)*, 38, 65-69.

Huang, X. (2014). *Exploring the mysteries of personality*. Beijing: Commercial Press.

Huang, X., & Zheng, Y. (2005). *Contemporary Chinese youth values research*. Beijing: People' s Education Press.

Guo, Y. (2018). *Outline of personality psychology*. Beijing: Educational Science Publishing House.

Jiang, C. (2018). Social recognition of the rationality and morality of core values. *Social Sciences in China*, 4, 4-23.

Li, L., & Huang, X. (2013). Neural mechanisms of values: An alternative research perspective. *Advances in Psychological Science*, 21(8), 1400-1408.

Yang, S., Huang, X., & Fu, Y. (2012). Medial prefrontal cortex—the neural basis of "self" . *Advances in Psychological Science*, 20(6), 853-859.

Yu, L., & Wang, L. (2013). Research on college students' implicit identification with socialist core values. *Journal of Southwest University (Social Sciences Edition)*, 39(5), 86-93.

Arieli, S., Grant, A. M., & Sagiv, L. (2014). Convincing yourself to care about others: An intervention for enhancing benevolence values. *Journal of Personality*, 82(1), 15-24.

Atran, S., & Ginges, J. (2012). Religious and sacred imperatives in human conflict. *Science*, 336(6083), 855-857.

Atran, S., & Ginges, J. (2015). Devoted actors and the moral foundations of intractable intergroup conflict. In J. Decety & T. Wheatley (Eds.), *The moral brain: A multidisciplinary perspective* (pp. 69-85). Boston Review.

Araujo, H. F., Kaplan, J., Damasio, H., & Damasio, A. (2014). Involvement of cortical midline structures in the processing of autobiographical information. *PeerJ*, 2, e481.

Bardi, A., Buchanan, K. E., Goodwin, R., Slabu, L., & Robinson, M. (2014). Value stability and change during self-chosen life transitions: Self-selection versus socialization effects. *Journal of Personality and Social Psychology*, 106(1), 131-147.

Bardi, A., & Goodwin, R. (2011). The dual route to value change: Individual processes and cultural moderators. *Journal of Cross-cultural Psychology*, 42(2),

271–287.

Baron, J. (2009). Belief overkill in political judgments. *Ssrn Electronic Journal*, 29(4), 368–378.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage*, 76, 412–427.

Berns, G. S., Bell, E., Capra, C. M., Prietula, M. J., Moore, S., Anderson, B., ··· Atran, S. (2012). The price of your soul: Neural evidence for the non-utilitarian representation of sacred values. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 754–762.

Bhanji, J. P., Beer, J. S., & Bunge, S. A. (2010). Taking a gamble or playing by the rules: Dissociable prefrontal systems implicated in probabilistic versus deterministic rule-based decisions. *NeuroImage*, 49(2), 1810–1819.

Boer, D., & Boehnke, K. (2015). What are values? Where do they come from? A developmental perspective. In D. Sander & T. Brosch (Eds.), *Handbook of value* (pp. 129–154). Oxford: Oxford University Press.

Brosch, T., Coppin, G., Scherer, K. R., Schwartz, S., & Sander, D. (2011). Generating value(s): Psychological value hierarchies reflect context-dependent sensitivity of the reward system. *Social Neuroscience*, 6(2), 198–208.

Brosch, T., Coppin, G., Schwartz, S., & Sander, D. (2012). The importance of actions and the worth of an object: Dissociable neural systems representing core value and economic value. *Social Cognitive and Affective Neuroscience*, 7(5), 497–505.

Brosch, T., & Sander, D. (2013). Neurocognitive mechanisms underlying value-based decision-making: From core values to economic value. *Frontiers in Human Neuroscience*, 7, 398.

Buhle, J. T., Silvers, J. A., Wager, T. D., Lopez, R., Onyemekwu, C., Kober, H., ···& Ochsner, K. N. (2014). Cognitive reappraisal of emotion: A meta-analysis of human neuroimaging studies. *Cerebral Cortex*, 24(11), 2981–2990.

Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience*, 7(4), 367–379.

Cooper, N., Bassett, D. S., & Falk, E. B. (2017). Coherent activity between brain regions that code for value is linked to the malleability of human behavior. *Scientific Reports*, 7, 43250.

Dogan, A., Morishima, Y., Heise, F., Tanner, C., Gibson, R., Wagner, A. F., & Tobler, P. N. (2016). Prefrontal connections express individual differences in intrinsic resistance to trading off honesty values against economic benefits. *Scientific Reports*, 6, 33263.

Döring, A. K., & Hillbrink, A. (2015). Brief report: Into the wild? How a film can change adolescents' values. *Journal of Adolescence*, 40, 78–82.

Duc, C., Hanselmann, M., Boesiger, P., & Tanner, C. (2013). Sacred values: Trade-off type matters. *Journal of Neuroscience, Psychology, and Economics*, 6(4), 252.

Ehret, P. J. (2019). *Inclusion of the environment in the self: Linking values to pro-environmental behaviors* (Unpublished doctoral dissertation). University of California, Santa Barbara.

Epton, T., Harris, P. R., Kane, R., van Koningsbruggen, G. M., & Sheeran, P. (2015). The impact of self-affirmation on health-behavior change: A meta-analysis. *Health Psychology*, 34(3), 187.

Falk, E. B., O' Donnell, M. B., Cascio, C. N., Tinney, F., Kang, Y., Lieberman, M. D., ···& Strecher, V. J. (2015). Self-affirmation alters the brain' s response to health messages and subsequent behavior change. *Proceedings of the National Academy of Sciences*, 112(7), 1977–1982.

Falk, E., & Scholz, C. (2018). Persuasion, influence, and value: Perspectives from communication and social neuroscience. *Annual Review of Psychology*, 69, 329–356.

Fransson, P., & Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *Neuroimage*, 42(3), 1178–1184.

Gardner, W. L., Gabriel, S., & Lee, A. Y. (1999). "I" value freedom, but "We" value relationships: Self-construal priming mirrors cultural differences in judgment. *Psychological Science*, 10, 321–326.

Gibson, R., Tanner, C., Wagner, A. F., Brosch, T., & Sander, D. (2016). Protected values and economic decision-making. In *Handbook of Value: Perspectives from Economics, Neuroscience, Philosophy, Psychology and Sociology* (p. 223).

Grant, A., & Sagiv, L. (2009). Convincing yourself to change your values: Self persuasion can increase the importance of benevolence values. Manuscript submitted for publication.

Hamid, N., Pretus, C., Atran, S., Crockett, M. J., Ginges, J., Sheikh, H., ··· & Vilarroya, O. (2019). Neuroimaging "will to fight" for sacred values: An empirical case study with supporters of an Al Qaeda associate. *Royal Society Open Science*, 6(6), 181585.

Hitlin, S. (2003). Values as the core of personal identity: Drawing links between two theories of self. *Social Psychology Quarterly*, 118–137.

Hitlin, S., & Civettini, N. (2017). The situated durability of values. *Advances in Group Processes*, 34(August 2017), 175–198.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: Consensus and controversy. *Neuron*, 63(6), 733–745.

Kaplan, J. T., Gimbel, S. I., Dehghani, M., Immordino-Yang, M. H., Sagae, K., Wong, J. D., ⋯& Damasio, A. (2016). Processing narratives concerning protected values: A cross-cultural investigation of neural correlates. *Cerebral Cortex*, 27(2), 1428–1438.

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology*, 22(6), 1027–1038.

Lönnqvist, J. E., Jasinskaja-Lahti, I., & Verkasalo, M. (2011). Personal values before and after migration: A longitudinal case study on value change in ingrian-finnish migrants. *Social Psychological and Personality Science*, 2(6), 584–591.

Lönnqvist, J. E., Jasinskaja-Lahti, I., & Verkasalo, M. (2013). Rebound effect in personal values: Ingrian Finnish migrants' values two years after migration. *Journal of Cross-Cultural Psychology*, 44(7), 1122–1126.

Lönnqvist, J. E., Leikas, S., & Verkasalo, M. (2018). Value change in men and women entering parenthood: New mothers' value priorities shift towards Conservation values. *Personality and Individual Differences*, 120, 47–51.

Lubinski, D., Schmidt, D. B., & Benbow, C. P. (1996). A 20-year stability analysis of the study of values for intellectually gifted individuals from adolescence to adulthood. *Journal of Applied Psychology*, 81(4), 443.

Mars, R. B., Sallet, J., Schüffelgen, U., Jbabdi, S., Toni, I., & Rushworth, M. F. (2012). Connectivity-based subdivisions of the human right "temporoparietal junction area" : Evidence for different areas participating in different cortical networks. *Cerebral Cortex*, 22(8), 1894–1903.

McNamee, D., Rangel, A., & O' doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience*, 16(4), 479.

Milfont, T. L., Milojev, P., & Sibley, C. G. (2016). Values stability and change in adulthood: A 3-year longitudinal study of rank-order stability and mean-level differences. *Personality and Social Psychology Bulletin*, 42(5), 572–588.

Pincus, M., LaViers, L., Prietula, M. J., & Berns, G. (2014). The conforming brain and deontological resolve. *PloS one*, 9(8), e106061.

Pretus, C., Hamid, N., Sheikh, H., Gomez, A., Ginges, J., Tobena, A., ⋯& Atran, S. (2019). Ventromedial and dorsolateral prefrontal interactions underlie will to fight and die for a cause. *Social Cognitive and Affective Neuroscience*, 14(6), 569–577.

Rokeach, M. (1973). *The nature of human values*. New York: Free press.

Rudnev, M. (2014). Value adaptation among Intra-European migrants: Role of country of birth and country of residence. *Journal of Cross-Cultural Psychology*,

45(10), 1626–1642.

Sandal, G. M., & Bye, H. H. (2015). Value diversity and crew relationships during a simulated space flight to Mars. *Acta Astronautica*, 114, 164–173.

Sandal, G. M., Bye, H. H., & Van De Vijver, F. J. R. (2011). Personal values and crew compatibility: Results from a 105 days simulated space mission. *Acta Astronautica*, 69(3-4), 141–149.

Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, 2(1), 1–20.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in Experimental Social Psychology* (Vol. 25, pp. 1–65). Academic Press.

Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology*, 32(5), 519–542.

Schwartz, S. H., & Rubel, T. (2005). Sex differences in value priorities: Cross-cultural and multimethod studies. *Journal of Personality and Social Psychology*, 89(6), 1010.

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences*, 109(42), 17058–17062.

Sheikh, H., Ginges, J., Coman, A., & Atran, S. (2012). Religion, group threat and sacred values. *Judgment and Decision Making*, 7(2), 110.

Sheldon, K. M. (2005). Positive value change during college: Normative trends and individual differences. *Journal of Research in Personality*, 39(2), 209–223.

Sherman, D. K., & Hartson, K. A. (2011). Reconciling self-protection with self-improvement. *Handbook of Self-enhancement and Self-protection*, 128.

Souza, M. J., Donohue, S. E., & Bunge, S. A. (2009). Controlled retrieval and selection of action-relevant knowledge mediated by partially overlapping regions in left ventrolateral prefrontal cortex. *Neuroimage*, 46(1), 299–307.

Sundberg, R. (2016). Value Stability and Change in an ISAF Contingent. *Journal of Personality*, 84(1), 91–101.

Sweeney, A. M., & Moyer, A. (2015). Self-affirmation and responses to health messages: A meta-analysis on intentions and behavior. *Health Psychology*, 34(2), 149.

Tao, L., & Au, W. T. (2014). Values, self and other-regarding behavior in the dictator game. *Rationality and Society*, 26(1), 46–72.

Verplanken, B., & Holland, R. W. (2002). Motivated decision making: Effects of activation and self-centrality of values on choices and behavior. *Journal of Personality and Social Psychology*, 82(3), 434.

Vezich, I. S., Katzman, P. L., Ames, D. L., Falk, E. B., & Lieberman, M. D. (2017). Modulating the neural bases of persuasion: Why/how, gain/loss, and users/non-users. *Social Cognitive and Affective Neuroscience*, 12(2), 283–297.

Wakslak, C., & Trope, Y. (2009). The effect of construal level on subjective probability estimates. *Psychological Science*, 20(1), 52–58.

Wei, T., Liang, X., He, Y., Zang, Y., Han, Z., Caramazza, A., & Bi, Y. (2012). Predicting conceptual processing capacity from spontaneous neuronal activity of the left middle temporal gyrus. *Journal of Neuroscience*, 32(2), 481–489.

Yomogida, Y., Matsumoto, M., Aoki, R., Sugiura, A., Phillips, A. N., & Matsumoto, K. (2017). The neural basis of changing social norms through persuasion. *Scientific Reports*, 7(1), 16295.

Young, L., Camprodon, J. A., Hauser, M., Pascual-Leone, A., & Saxe, R. (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*, 107(15), 6753–6758.

Zacharopoulos, G., Hanel, P. H., Lancaster, T. M., Ihssen, N., Drakesmith, M., Foley, S., ···& Linden, D. E. (2017). Nonlinear associations between human values and neuroanatomy. *Social Neuroscience*, 12(6), 673–684.

Zacharopoulos, G., Lancaster, T. M., Bracht, T., Ihssen, N., Maio, G. R., & Linden, D. E. (2016). A hedonism hub in the human brain. *Cerebral Cortex*, 26(10), 3921–3927.

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences*, 104(15), 6430-6435.

Zahn, R., Moll, J., Paiva, M., Garrido, G., Krueger, F., Huey, E. D., & Grafman, J. (2009). The neural basis of human social values: Evidence from functional MRI. *Cerebral Cortex*, 19, 276–283.

Note: Figure translations are in progress. See original paper for figures.