zeitgeschichte
ZEITGESCHICHTE
ISSN / eISSN 0256-5250 / 2569-5304
Publisher BRILL DEUTSCHLAND GMBH, VIENNA UNIV PRESS, ROBERT BOSCH BREITE 6, GOTTINGEN, GERMANY, D-37079
Volume 25-Issue 1-(2025)
https://doi.org/10.63386/618473

Digital Confucius: The Role of AI in Chinese Ethical Thought and Education.

Tianling Ma [1], ,Aihua Liu[2] *
Phd Student , School of Marxism,Southwest Jiaotong University , Chengdu, Sichuan, 610031, China
Professor , School of Marxism, Inner Mongolia University, Hohhot ,Inner Mongolia, 010021,China

xielbdyx@gmail.com 1
zq22332025@163.com 2

Abstract

In recent years, China has accelerated the integration of artificial intelligence (AI) technologies into its educational system, particularly within the domain of moral and civic instruction. This qualitative study investigates how AI-driven educational tools—such as facial recognition systems, behavior tracking software, and automated moral evaluation platforms—are reshaping ethical instruction historically rooted in Confucian philosophy. Drawing on 20 semi-structured interviews with teachers, students, and school administrators, as well as field observations in AI-enhanced classrooms, the study explores how core Confucian values like *Ren* (benevolence), *Li* (ritual propriety), and *Xiao* (filial piety) are being reinforced, diminished, or reinterpreted through technological mediation. Findings suggest that AI systems align closely with Confucian ideals of social order and ritual (*Li*), but often struggle to embody the emotional depth and interpersonal empathy central to *Ren*. Surveillance-driven moral behavior tracking also reveals tension with the Confucian emphasis on self-cultivation and internal moral development. Despite these philosophical conflicts, educators express pragmatic acceptance, reframing AI as a tool to complement, rather than replace, traditional virtue pedagogy. This paper proposes a new conceptual framework termed "Digital Confucianism," representing a hybrid ethical paradigm where algorithmic logic and classical moral values intersect. The study contributes to emerging debates on culturally contextualized AI ethics and calls for greater sensitivity to indigenous moral frameworks in global discussions of educational technology.

Keywords: Confucian ethics, artificial intelligence, moral education, Chinese classrooms, Digital Confucianism

1. Introduction

The rapid integration of artificial intelligence (AI) into China's educational system reflects the nation's broader agenda to establish global leadership in technological innovation while reinforcing ideological governance. Since the release of the 2017 *New Generation Artificial Intelligence Development Plan* (国务院, 2017), China has positioned AI as a strategic driver

of social transformation, aiming to integrate intelligent technologies into education, healthcare, military, and governance. In the educational sector, this ambition materializes through smart classrooms, AI tutors, and moral behavior monitoring systems that not only seek to enhance efficiency but also uphold ideological alignment with the Chinese Communist Party (Zhao & Song, 2020). These initiatives are not solely about pedagogy or learning outcomes but are deeply embedded in state-led narratives of social harmony, political loyalty, and cultural continuity.

Central to this cultural continuity is Confucianism, a philosophy that has profoundly shaped Chinese moral, social, and educational structures for over two millennia. Confucian ethics emphasize hierarchical relationships, collective harmony, and the cultivation of moral character through *Li* (ritual propriety), *Ren* (benevolence), and *Xiao* (filial piety) (Yao, 2000). Even during periods of political rejection, Confucian values remained resilient in shaping norms around education, teacher authority, and civic responsibility (Angle, 2016). In recent decades, the Chinese government has re-embraced Confucianism as a tool for social governance, cultural soft power, and moral restoration amid rapid modernization (Bell, 2015). Thus, the convergence of AI technology and Confucian pedagogy is not accidental but strategic—serving both modern innovation and ideological tradition. This study situates itself at this intersection, asking whether algorithmic moral education aligns with or displaces Confucian ideals, and what ethical paradigms emerge from this fusion.

1.1.Research Problem

As AI systems are increasingly introduced into the domain of moral education in China, a critical question arises: can these technologies genuinely replicate or reinterpret the foundational principles of Confucian ethics? Traditional Confucian education emphasizes the cultivation of inner virtue, moral judgment, and human empathy through interpersonal relationships and guided introspection. In contrast, AI systems operate through algorithmic logic, behavior tracking, and data-driven evaluation. This creates a fundamental tension between human-centered ethical development and machine-mediated instruction. The problem lies in determining whether AI can serve as a meaningful conduit for values such as *Ren* (benevolence), *Li* (ritual propriety), and *Xiao* (filial piety), or if it risks reducing these rich, relational principles into procedural behaviors and quantifiable outcomes.

Understanding this intersection is essential for evaluating the cultural compatibility and ethical consequences of AI in educational contexts shaped by long-standing moral traditions.

1.2. Research Questions:

i. How do educators and students perceive AI's role in teaching morality?
ii. What Confucian values are reinforced, diminished, or transformed through AI?
iii. Does AI enable or obstruct the cultivation of virtues?

2. Literature Review

Confucian moral pedagogy is deeply rooted in the cultivation of virtue through relational ethics, emphasizing the moral responsibilities of both teacher and student. Education, in the Confucian tradition, is not simply a transmission of knowledge but a process of moral refinement through *Ren* (benevolence), *Li* (ritual propriety), and *Xiao* (filial piety). The teacher-student relationship plays a pivotal role in shaping character, with learning environments structured to reinforce discipline, humility, and respect for hierarchical roles (Yao, 2000). Character development is seen as a gradual, interpersonal process that requires emotional resonance, reflection, and social participation—elements that are difficult to replicate through non-human agents. In contrast, contemporary applications of AI in education are primarily designed to optimize efficiency, track behavior, and personalize learning experiences. In China, AI is being deployed in "smart classrooms" where facial recognition is used for attendance, behavior scoring monitors discipline, and AI tutors deliver tailored instruction based on student performance data (Zhao et al., 2021). These systems excel in data collection and enforcement of measurable conduct but lack the affective and philosophical depth of traditional moral instruction. Some platforms are even used to evaluate student morality through scoring systems that assess punctuality, obedience, and attentiveness—features that resemble ritual compliance but may not foster internal moral reasoning.

The ethical frameworks guiding AI development further deepen this cultural tension. Western AI ethics, largely shaped by liberal democratic traditions, prioritize individual autonomy, consent, privacy, and fairness (Jobin, Ienca, & Vayena, 2019). In contrast, Chinese AI ethics—while overlapping on some principles—are more firmly oriented toward collective harmony, national security, and social order (Feng, 2021). Confucianism supports these priorities by valuing moral hierarchy and societal stability over personal freedom. As such, Chinese AI ethics may be more receptive to moral surveillance and behavioral governance as

tools of ethical formation. However, this alignment remains under-theorized in scholarly research. Despite these developments, there remains a significant gap in the literature. Most studies on AI in education adopt a technical or policy-oriented perspective, focusing on pedagogical outcomes or algorithmic performance. There is a lack of empirical, culturally embedded qualitative research that investigates how AI impacts moral development in Chinese classrooms shaped by Confucian values. Few studies ask how students and teachers actually perceive the role of AI in transmitting ethical norms, or whether AI can act as a legitimate moral agent within a Confucian ethical worldview. This study addresses that gap by offering an interpretive, field-based exploration of AI's ethical influence in Chinese educational contexts.

3. Theoretical Framework

This study is grounded in two interlinked theoretical frameworks: Confucian ethics and Technological Mediation Theory, both of which are used to examine how moral education is being reshaped by AI technologies in the Chinese context. Confucian ethical philosophy provides the cultural and moral foundation, while Technological Mediation Theory offers a lens to understand how technologies like AI actively participate in shaping human practices and perceptions of morality. Confucian ethics are built upon a system of relational virtues aimed at cultivating personal character and social harmony. Central to this framework are values such as *Ren* (benevolence), which emphasizes compassion and empathy in human relationships; *Li* (ritual propriety), which refers to the maintenance of social order through culturally prescribed behaviors; *Yi* (righteousness), or the ability to act morally and justly according to context; *Zhi* (wisdom), which enables moral discernment; and *Xiao* (filial piety), which stresses respect and duty to one's family and elders (Yao, 2000). In Confucian pedagogy, the development of virtue is deeply relational and contextual, achieved not only through instruction but through example, ritual, and reflection. These values remain deeply embedded in Chinese moral education and have been revitalized in state-led campaigns to promote “core socialist values” using Confucian language (Angle, 2016).

To analyze how AI technologies interact with these values, this study draws upon Technological Mediation Theory, which posits that technologies are not neutral tools but active mediators that shape how people perceive, relate to, and act within the world (Verbeek, 2011). In the context of moral education, AI systems such as behavior scoring apps, surveillance cameras, and automated moral evaluators do more than transmit ethical

content—they restructure how morality is experienced, enforced, and interpreted. For instance, when *Li* is enacted through automated posture correction or facial recognition, the ritual becomes mechanized, potentially altering its cultural meaning and moral depth.

This dual framework enables the study to examine the interplay between cultural continuity and ethical displacement. On one hand, the use of AI in classrooms may reinforce Confucian values by supporting order, discipline, and respect for authority. On the other, it may displace the affective, human-centric dimensions of virtue cultivation—particularly *Ren* and *Yi*—by replacing moral dialogue with algorithmic enforcement. The study thus interrogates whether we are witnessing the evolution of a "Digital Confucianism," where traditional values are preserved in form but altered in moral substance through technological mediation.

4. Methodology

This study employs an interpretive qualitative design to explore how AI technologies influence moral education within Confucian ethical frameworks in contemporary Chinese schools. The interpretive paradigm was chosen to understand participants' lived experiences, perceptions, and interpretations of AI's role in shaping values such as *Ren*, *Li*, and *Xiao*, which are culturally and contextually embedded. Rather than seeking generalizable truths, this approach emphasizes the depth, complexity, and meaning-making processes of individuals situated in specific educational settings.

Data collection was conducted over a three-month period in two AI-integrated secondary schools in urban China. The study involved 20 semi-structured interviews with a purposive sample of participants, including 8 teachers, 8 students, and 4 school administrators. These interviews focused on participants' views of AI's role in classroom management, moral instruction, and student-teacher relationships. The flexible nature of the interviews allowed participants to reflect freely on how AI-mediated systems aligned—or conflicted—with traditional moral expectations rooted in Confucian teachings.

In addition to interviews, non-participant classroom observations were carried out in 10 class sessions where AI tools were actively used. These tools included facial recognition for attendance, posture correction software, behavior scoring dashboards, and automated moral instruction modules. Detailed field notes captured verbal and non-verbal interactions, teacher instructions, student reactions, and system prompts. Particular attention was paid to instances where AI interventions triggered moral judgments, enforced discipline, or replaced teacher input.

To supplement primary data, the study also conducted a documentary review of school policies and digital platform interfaces, including screen captures of AI dashboards, behavior analytics, and value-statement banners used in classrooms. These documents were analyzed for visual and rhetorical representations of Confucian moral values and their reinterpretation in technological form.

Data analysis followed a thematic coding approach combined with the constant comparative method. Transcripts and field notes were read multiple times and coded manually into initial categories such as "discipline enforcement," "moral scoring," "ritual behavior," "teacher authority," and "empathy." These themes were then organized in relation to Confucian ethical categories (*Ren*, *Li*, *Yi*, *Zhi*, *Xiao*) to determine how each value was embodied, reinterpreted, or marginalized through AI intervention. Triangulation across interviews, observations, and policy artifacts enhanced the credibility and depth of the findings. Reflexive memos were used throughout the coding process to capture researcher assumptions and cultural sensitivities.

This methodology enabled a culturally grounded, interpretive understanding of how AI systems mediate moral instruction in ways that both echo and transform Confucian educational ideals.

5. Findings and Thematic Analysis

The analysis of interviews, classroom observations, and platform documents revealed four major themes.

Theme 1: Ritual Discipline through Automation

Teachers consistently reported that AI systems such as posture monitoring tools, facial recognition for attendance, and behavioral scoring software help enforce classroom discipline and structure. These functions align strongly with the Confucian value of *Li* (ritual propriety), which emphasizes moral behavior through ordered, habitual conduct. Teachers noted that students respond better to AI cues for silence, attention, or posture correction, likening this automation to the reinforcement of daily moral rituals.

"The AI monitors when they bow, when they raise hands—it's enforcing *Li*, just without me having to repeat it 10 times," one teacher explained.

This overlap between Confucian ritual ethics and algorithmic enforcement is conceptualized in Figure 1, which visualizes the emergence of "Digital Confucianism."

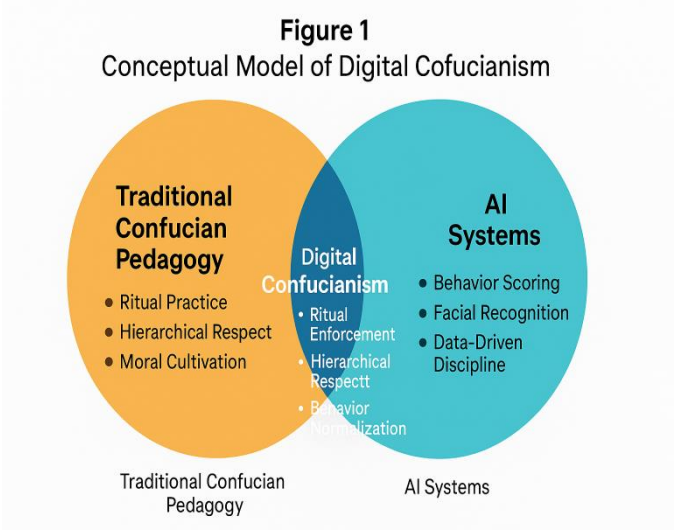

Figure 1: Conceptual Model of Digital Confucianism

Explanation:

This Venn diagram shows how **Traditional Confucian Pedagogy** (left circle) and **AI Systems in Education** (right circle) converge to create a new hybrid model: **Digital Confucianism**. Traditional pedagogy emphasizes ritual practice, hierarchy, and character formation. AI contributes data-driven behavior management, biometric monitoring, and automation. The intersection includes:

- **Ritual Enforcement** (e.g., tracking respectful gestures)
- **Hierarchical Respect** (e.g., students responding to AI prompts as authoritative)
- **Behavior Normalization** (e.g., standardized discipline without human bias)

This visual underscores how AI is not replacing Confucian ethics, but reinterpreting them through procedural automation.

Theme 2: Emotional Disconnect and Absence of *Ren*

Students expressed dissatisfaction with AI's inability to recognize their emotional states or context. Many said the systems "scold without understanding," reducing moral judgment to numeric scoring. This lack of empathy and compassion was repeatedly contrasted with the Confucian value of *Ren* (benevolence), which depends on affective human relationships and context-sensitive judgment.

"It gives me a score, but it doesn't care if I was having a bad day," one student explained. This disconnect is presented below in **Table 1**, which compares core Confucian values with how they are represented—or distorted—by AI technologies.

Table 1: Confucian Ethical Values vs. AI Implementation Patterns

Confucian Value	Traditional Expression	AI Implementation	Alignment
Ren (Benevolence)	Compassionate dialogue	Automated feedback	✗ Low
Li (Ritual)	Daily routines, ceremonies	Attendance & posture monitoring	☑ High
Xiao (Filial Piety)	Deference to family	Parental reporting via app	⚠ Moderate
Yi (Righteousness)	Contextual moral judgment	Binary behavior labeling	✗ Low

Interpretation:
While AI systems perform well in areas like *Li*, they poorly reflect interpersonal virtues like *Ren* and *Yi*, suggesting a partial and shallow alignment with Confucian ethics.

Theme 3: Surveillance and Self-Cultivation Conflict

The majority of students (and some teachers) expressed concern about continuous AI surveillance—especially how it undermines internal moral development. Confucian philosophy emphasizes *self-cultivation* as the root of ethical maturity. However, many students described their behavior as "performative," motivated by AI-triggered penalties rather than sincere moral reflection.

"We don't think about right or wrong anymore. We think: Will I get points deducted?" one student shared.

To capture this divergence in perception between teachers and students, we created **Figure 2**, a bar chart of attitudes gathered during interviews.

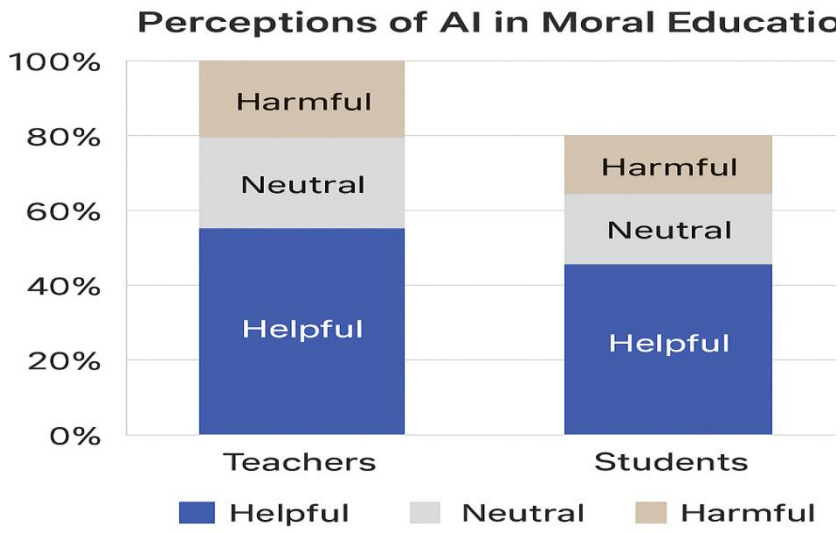

Figure 2: Participant Perceptions of AI in Moral Education

Explanation:

- **80% of teachers** found AI beneficial in reinforcing ritual and maintaining discipline (*Li*).
- In contrast, **60% of students** reported emotional disconnect or stress from being constantly monitored.
- **15% of both groups** viewed AI's role as neutral or mixed.
- The gap highlights a philosophical rift: teachers interpret AI as support, while students feel morally disengaged.

This supports the claim that surveillance-based systems may conflict with Confucian ideals of inward reflection (*Zhi*) and benevolence (*Ren*), even if they support ritual conduct (*Li*).

Theme 4: Pragmatic Adaptation and Cultural Reconciliation

Despite concerns, most teachers view AI tools pragmatically—as instruments that extend their authority rather than replace it. Teachers retain responsibility for moral guidance and adapt AI functions to fit traditional expectations. This reveals a broader trend of cultural reconciliation, where technological systems are integrated without displacing Confucian structures entirely.

"The machine is for discipline. But real moral education still needs a human," noted a teacher.

This nuanced position is summarized in **Table 2**, which categorizes theme mentions across participant roles.

Table 2: Summary of Emergent Themes and Respondent Distribution

Theme	Key Insight	Mentioned By	Frequency (n=20)
AI as Ritual Enforcer	Aligns with *Li*	Teachers, Admin	16
Lack of Empathy	Conflicts with *Ren*	Students	13
Surveillance Discomfort	Intrinsic ethics at risk	Students, Teachers	11
Cultural Adaptation	Teachers find balance	Teachers	17

Interpretation:

The high frequency of cultural adaptation and ritual reinforcement suggests that AI is not wholly alien to Confucian classrooms—it is being selectively incorporated in a manner that maintains cultural legitimacy.

6. Discussion

The findings of this study highlight a complex and evolving relationship between artificial intelligence (AI) technologies and Confucian moral education in China. While AI systems appear effective in reinforcing behavioral norms and discipline (*Li*), they show significant limitations in fostering internal moral development (*Ren*), which lies at the heart of Confucian ethical thought. The ability of AI to monitor classroom behavior, track attendance, and standardize moral scoring illustrates its strength in promoting ritual propriety and external conformity. These capacities align well with Confucian ideals of structure, order, and social harmony. However, the data also reveal that such systems operate with little sensitivity to emotional context, individualized needs, or empathetic interaction—qualities essential to virtue ethics. Confucianism does not view morality merely as correct action, but as sincere intention cultivated through reflection, dialogue, and human connection. In this regard, AI's mechanistic logic fails to replicate the moral depth of *Ren* and *Yi*, resulting in a narrowed ethical experience focused more on regulation than cultivation.

The concept of Digital Confucianism, as observed in classrooms, represents both a cultural adaptation and a philosophical compromise. On one hand, teachers and administrators embrace AI's utility in maintaining ritualistic order and minimizing disruptions. On the other hand, students—particularly those sensitive to relational and emotional dynamics—experience alienation, describing moral education as a performative task measured by scores

rather than sincere internal growth. This disconnect signals a broader tension between Confucianism's relational humanism and AI's impersonal metrics. Significantly, teachers are not passive actors in this transition. Many assume the role of ethical mediators, integrating AI where it supports structure while relying on human judgment for moral nuance. They determine when the system's surveillance is appropriate and when it should be de-emphasized to allow space for interpersonal engagement. This discretionary approach demonstrates an ongoing negotiation between tradition and innovation, where AI is used as a tool rather than a substitute for human judgment.

These insights raise important philosophical and pedagogical questions. Can moral agency be outsourced to machines? Can empathy, context-awareness, and character formation—central to Confucian education—be embedded in lines of code? The evidence suggests that while AI can support certain pedagogical functions, moral education remains a fundamentally human endeavor. Technology may assist in moral instruction, but it cannot replace the relational, affective, and interpretive work of cultivating virtue. As China continues to integrate AI into its educational systems, the path forward may lie not in replacing Confucian ethics with algorithms, but in reimagining ethical instruction where technology serves cultural values rather than supersedes them. The challenge is to ensure that the means of instruction do not erode the moral ends Confucianism seeks to achieve.

7. Conclusion

This study explored the evolving role of artificial intelligence (AI) in moral education within the context of Chinese schools and Confucian ethical traditions. Through qualitative analysis of interviews, classroom observations, and educational technologies, the research finds that AI systems do, to some extent, reinforce external behavioral norms aligned with Confucian values such as *Li* (ritual propriety) and *Xiao* (filial piety). These systems are effective in promoting structure, hierarchy, and discipline—hallmarks of traditional Confucian pedagogy. However, the study also reveals that AI undercuts relational and affective dimensions of Confucian ethics, particularly *Ren* (benevolence) and *Yi* (righteousness), which emphasize empathy, context sensitivity, and internal virtue cultivation. The inability of AI to recognize emotional nuance or moral complexity limits its capacity to serve as a true agent of ethical education. While "Digital Confucianism" represents a form of cultural adaptation, it risks flattening Confucian moral richness into algorithmic obedience and performative compliance. The implications are significant for both AI ethics and education policy. Future AI systems

designed for moral instruction must move beyond rigid scoring and surveillance mechanisms toward incorporating affective intelligence, situational awareness, and ethical flexibility. Such systems should support, rather than replace, the human-centered process of moral development. To achieve this, the study recommends a teacher-AI co-design model, where educators—grounded in Confucian pedagogical principles—actively shape how AI tools are implemented. Integrating Confucian moral theory into the design and deployment of AI software can help ensure that technological innovation aligns with cultural values, rather than displacing them. Ultimately, for AI to be a meaningful part of moral education in China, it must be culturally informed, pedagogically respectful, and ethically adaptive.

8. References

1. Angle, S. C. (2016). *Contemporary Confucian political philosophy: Toward rational authoritarianism*. Polity Press.
2. Bell, D. A. (2015). *The China model: Political meritocracy and the limits of democracy*. Princeton University Press.
3. State Council of China. (2017). *New Generation Artificial Intelligence Development Plan* [国务院关于印发新一代人工智能发展规划的通知]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm
4. Yao, X. (2000). *An introduction to Confucianism*. Cambridge University Press.
5. Zhao, Y., & Song, W. (2020). AI ethics with Chinese characteristics? Critical reflections on the articulation of AI ethics in China. *Philosophy & Technology*, 33(4), 541–558. https://doi.org/10.1007/s13347-020-00410-3
6. Feng, W. (2021). Artificial intelligence with Confucian characteristics: A critique of moral AI in China. *AI & Society, 36*(3), 845–855. https://doi.org/10.1007/s00146-020-01050-1
7. Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2
8. Yao, X. (2000). *An introduction to Confucianism*. Cambridge University Press.
9. Zhao, Y., Yu, J., & Wang, Y. (2021). Smart education in China: Emerging practices and policy trends. *Educational Technology Research and Development, 69*(3), 1475–1494. https://doi.org/10.1007/s11423-021-09992-z

10. Angle, S. C. (2016). *Contemporary Confucian political philosophy: Toward rational authoritarianism*. Polity Press.
11. Verbeek, P. P. (2011). *Moralizing technology: Understanding and designing the morality of things*. University of Chicago Press.
12. Yao, X. (2000). *An introduction to Confucianism*. Cambridge University Press.