ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

Who Defines a “Good Response"? Cultural Adaptation in Generative AI Interactions in Family Contexts

Siying Li[1, 2], Chen Gao[1]

[1]Graduate School of Global Business, Kyonggi University, Suwon 16227, South Korea

[2]School of Management, University of Science and Technology Liaoning, Liaoning 114000, China

Abstract

As generative artificial intelligence moves more deeply into family life, it may influence how children seek information, express emotion, and form value judgments. The normative problem examined here is therefore not simply whether an AI response appears friendly or reasonable. A further question is whether the same response carries an equivalent educational meaning across different ethical contexts of family life. To address that question, this article treats ren (benevolence), yi (righteousness), li (propriety), zhi (practical wisdom), xin (trustworthiness), xiao (filiality), he (harmony), and chi (moral shame) as analytical lenses. Comparative cultural analysis, conceptual analysis, normative reconstruction, and eight thought experiments are then brought together to consider two sides of AI responses: the values they protect and the relational conditions they may leave insufficiently examined. Five dimensions organize the tensions identified through this analysis: relationality, regulation, discernment, truthfulness, and motivation. On that basis, the article proposes design principles concerning relational sensitivity, age appropriateness, honest feedback, and responsibility-oriented reflection. The framework should, however, be read within its stated limits. It is intended as a contestable and revisable theoretical construction, rather than as empirical evidence about actual parental attitudes or child developmental outcomes.

Keywords

Generative artificial intelligence; Confucian ethics; parenting; cultural adaptation; human-computer interaction; cross-cultural research.

1. Introduction

1.1. Background and Problem Statement

Information seeking, learning, emotional expression, and decision-making are already supported by conversational systems based on large language models. Their spread has widened both the forms and the practical boundaries of human-computer interaction. Once the users are children, however, usefulness is only one criterion: age appropriateness and rights protection also become conditions of acceptable use. UNESCO's guidance emphasizes human-centered and age-appropriate educational applications and recommends age limits for children's independent use of generative AI (Miao & Holmes, 2023). UNICEF (2021), approaching the issue through child rights, likewise identifies safety, privacy, fairness, transparency, and well-being as necessary requirements. Within family contexts, the question is therefore not exhausted by whether AI can generate an answer. The route by which that answer enters parent-child communication, and the conditions under which it participates in children's socialization, also require examination.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

A friendly or reasonable appearance does not by itself establish a stable educational meaning. As AI enters children's family lives and processes of moral socialization, the same response may, at least in principle, be understood differently across cultural and family contexts. This is the second-order question addressed here; surface acceptability alone does not resolve it.

To describe generative AI responses as wholly value-neutral would overlook the influence of model training, preference optimization, safety rules, and product objectives on both tone and position. One relevant tendency is identified in existing research as "sycophancy": a model may accommodate a user's position at the expense of truthfulness or independent judgment. Experiments with open-ended tasks suggest that several AI assistants tend to align with users' views (Sharma et al., 2024; Wei et al., 2023). Yet those findings do not establish a fixed script across products. Accordingly, "initial affirmation, followed by softened disagreement and an open-ended question" is treated in this article as a possible response pattern, not as a universal product property. Whether it appears, and with what intensity, remains a question for particular models and versions.

Research on AI and children has so far been organized mainly around technological ethics, privacy protection, and cognitive development. Less sustained attention has been given to **cultural specificity**. The term is used here in a limited sense: the same AI interaction strategy may be evaluated differently across cultural settings. It is this possible variation in ethical perception and educational judgment, rather than a claim that cultures are internally uniform, that provides the theoretical point of departure for the present study.

1.2. Research Aim and Questions

The study does not set out to rank complete ethical traditions. Its aim is narrower: to construct an ethical framework capable of explaining cultural tensions in generative AI interactions within families. For that purpose, Confucian relational ethics is placed in comparison with educational discourses such as autonomy support and positive psychological adjustment. The comparison asks how differences in value priority may shape judgments about what should count as a "good response."

Three questions organize the inquiry:

1. Which forms of generative AI response may become matters of concern in family ethics, and which considerations lie at the center of those concerns?

2. How can the reasoning behind such concerns be reconstructed, and in what ways does that reasoning connect to Confucian categories of virtue?

3. How does the standard of a "good response" implicit in AI response strategies differ structurally from the standard associated with Confucian parenting contexts, and which value assumptions and cultural logics account for the difference?

1.3. Significance of the Study

Theoretical significance: One contribution lies in bringing a systematic set of Confucian ethical categories into human-computer interaction research, thereby providing a culturally grounded framework for examining adaptive AI design. A second lies in the treatment of generative AI as a possible third-party agent within child socialization, which extends discussion in cross-cultural psychology and the philosophy of education. The contribution is therefore not merely the addition of cultural examples to an existing model. In this qualified sense, AI cannot be treated as a neutral instrument alone, because assumptions in its responses may carry particular educational and cultural priorities.

Practical significance: A relational-ethical vocabulary allows family-oriented generative AI to be reviewed for consequences that factual correctness, immediate user satisfaction, or individual preference may not disclose. Its purpose is to generate questions for research and product development, not to replace studies with actual users or prescribe one parenting model

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

across families. Four design requirements follow provisionally: family disagreement must be recognized; explanatory depth, choice architecture, and adult involvement should vary with developmental stage; factual correction and substantive disagreement must remain possible within a supportive tone; and feedback about failure should connect causal analysis to concrete corrective action. Relational-context detection, age-differentiated responses, and calibration of facts and positions are possible directions. Another is a staged pathway from advice to negotiated dialogue and, where appropriate, to the involvement of a trusted adult. These proposals remain design principles; their technical implementation and effectiveness have not been demonstrated here.

1.4. Structure of the Article

The argument is developed across six sections. After this introduction, Section 2 reviews the literature and establishes the theoretical coordinates of the comparison. Section 3 explains the methodological combination of conceptual analysis, normative reconstruction, and thought experiments. The Confucian virtue framework is developed in Section 4 through eight analytical thought experiments; Section 5 then examines structural differences between relationally oriented and individually oriented educational discourses. Section 6 draws the theoretical claims together, proposes design principles, and specifies the conditions and boundaries of the framework's application.

2. Literature Review and Theoretical Framework

2.1. Existing Research on Children's Interactions with Generative AI

Research on children's interactions with generative AI can be read as three partly overlapping strands. Cognition and learning form one strand, including possible effects on information retrieval, problem solving, and critical thinking. Although this evidence base is expanding rapidly, stable causal conclusions specifically about primary-school children are not yet available. A second strand concerns interaction and trust, especially anthropomorphism, accommodating responses, and possible effects on user judgment. Studies of sycophancy show that large language models may accommodate users' positions (Sharma et al., 2024; Wei et al., 2023), but they do not establish what follows in children's family contexts. Ethics and governance supply the third strand, in which children's rights, privacy, safety, fairness, transparency, and age appropriateness are prominent (UNICEF, 2021; Miao & Holmes, 2023). Across the three strands, a culturally sensitive and systematic account of how response styles enter family relationships and moral socialization remains comparatively underdeveloped.

This literature is indispensable to the governance of children's AI use, but it does not settle the question pursued here. Might the educational meaning of one response strategy change when the ethical context of family life changes? Cultural intuition cannot answer that question, nor can the experience of particular families be treated as sufficient evidence. The inquiry therefore begins at the theoretical level. Its framework is meant to prepare questions for subsequent empirical testing, rather than to anticipate the answers.

2.2. Confucian Ethics and Child Socialization: Analytical Definitions of Core Categories

For the purposes of the present analysis, ren, yi, li, zhi, xin, xiao, he, and chi are used as interpretive lenses. Their assembly into an eight-part framework is specific to this study of parenting and should not be mistaken for a fixed classification inherited directly from the Confucian classics. Ren denotes care, drawing on the injunction to "love others" in the Yan Yuan chapter of the Analects. Yi refers to what ought to be done within a role and is informed by the statement that "the exemplary person understands what is right" in the Li Ren chapter. Li concerns interactional order and appropriate measure. Xin emphasizes the reliability of a

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

response, consistent with the statement in the Wei Zheng chapter that "If a person is not trustworthy, one does not know what can be made of that person." Xiao concerns respect and responsibility across generations (Wei Zheng), whereas he foregrounds the view that "in the practice of ritual propriety, harmony is most valuable" (Xue Er). Chi has a self-corrective meaning, expressed in the Doctrine of the Mean within the Book of Rites through the claim that "To know shame is close to courage." The study's account of zhi also draws on the value-ordering implication of the Great Learning: "only after knowing where to stop can one attain steadiness." These definitions extend classical resources for the analysis of AI interactions in family contexts. They remain analytical choices and are neither the sole nor exhaustive interpretations of the concepts (Wang, 2001; Yang, 2006).

Section 4.1 presents the working definition of each category and the tension associated with it. These categories are interpretive concepts reconstructed from Confucian ethical resources; they should not be read as empirical descriptions of the psychological structures of actual parents. Considered together, they form a relational-ethical framework in which individual judgment is situated within family roles, intergenerational responsibilities, and interpersonal ties. This way of locating the person relationally is broadly consonant with scholarship on the Confucian self and the differential mode of association (Hall & Ames, 1998; Fei, 1998; Hwang, 2009), although the framework developed here remains tailored to the present research question.

2.3. Comparative Points of Departure for Chinese and Western Ethical Frameworks

Before the three comparative dimensions are used, their status needs to be limited in two respects. They are ideal-typical axes constructed to explain differences in emphasis, not essentialist descriptions of Chinese and Western societies. They also leave room for substantial variation within cultural groups and among individuals, even though cross-cultural psychology has identified different relative weights assigned to autonomous and relational models of the self (Markus & Kitayama, 1991). A further qualification concerns self-determination theory: autonomy there does not mean detachment from relationships or refusal of rules, but volition and the internal endorsement of action (Deci & Ryan, 2000). What follows is thus a comparison of value priorities and modes of expression, not a simple opposition between internally uniform traditions.

Dialogue provides the first comparative axis. Where parent-child dialogue is understood mainly as a setting for expressing views, exercising thought, and maintaining personal boundaries, support for autonomous expression becomes one criterion for a "good response." Confucian parenting traditions may order attention differently by assigning greater initial weight to relational co-presence, tacit understanding, and the reinforcement of emotional bonds. From that standpoint, the judgment also asks whether harmonious co-presence is preserved. The two questions need not exclude one another. Their difference lies, more cautiously, in which consideration receives initial priority.

Second, the traditions differ in how virtue develops. Within some mainstream Western developmental theories, including self-determination theory, meaningful choice and empowerment are conditions through which autonomy can emerge. Confucian parenting assigns comparatively greater weight to instruction and internalization: a child first encounters the meaning and reasonableness of a rule, then acquires a fuller voice in discussing it. This is a difference of emphasis rather than a universal sequence for every family. It becomes especially salient when AI provides a young child with the language of empowerment while leaving the relevant reasons and developmental support underexplained.

How error and failure are interpreted forms the third axis. Some positive psychology interventions reframe failure as an opportunity for growth, in part to protect children's self-

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

efficacy. Confucian traditions can accept that developmental function and still attach a role-oriented meaning to failure. Under the latter interpretation, an assessment of ability does not exhaust the event; an unfulfilled role obligation may also be involved. For chi, the motivational function arises not from general self-condemnation, but from recognition of a shortcoming that remains open to correction.

3. Methodology and Theoretical Approach

3.1. Methodological Positioning

The methodology coordinates comparative cultural analysis, conceptual analysis, and normative reconstruction. Through comparative cultural analysis, value assumptions drawn from different traditions are juxtaposed (Shweder, 1991; Nisbett, 2003). Conceptual analysis clarifies the meanings and limits of autonomy, care, responsibility, and contextual appropriateness. Normative reconstruction then asks which value may be placed under strain when a response principle protects another. Eight thought experiments provide imagined cases in which conceptual conflicts can be exposed, principles tested for consistency, and counterexamples considered. Because they are imagined rather than observed, they cannot substitute for real-world evidence (Gendler, 2000). The analysis therefore works with ideal types while leaving open the internal diversity of both "Chinese families" and "Western education."

Placing key concepts from two cultural contexts in relation to one another is the central procedure of comparative cultural analysis (Shweder, 1991; Nisbett, 2003). This relational placement makes it possible to examine conceptual structure, value assumptions, and practical logic as connected patterns rather than as isolated surface contrasts. Comparison nevertheless requires movement between two perspectives. An emic perspective seeks to understand a tradition's rationality from within; an etic perspective makes cross-cultural comparison possible. The present argument depends on both, since neither is adequate on its own.

Conceptual analysis here begins where lexical definition ends. In the philosophy of education, one way of clarifying an educational judgment is to ask under which normative conditions a concept is appropriately used. Peters's (1967) account of the normative content of "education," for instance, shows that educational concepts already contain judgments about forms of development considered worth pursuing. Against that background, the present study examines how the semantic boundaries and normative assumptions of a "good response" shift across autonomy, care, responsibility, and contextual appropriateness. Searle's (1969) speech act theory plays a supporting role by treating speaking as action. Comforting, advising, correcting, and persuading can therefore differ not only as modes of transmitting information, but also in their educational effects. Peters supplies the basis for normative conceptual clarification; Searle helps distinguish response acts and the conditions under which their functions change.

3.2. Objects of Analysis and Source Materials

Three groups supply the materials for analysis. Research on children's rights, educational uses of generative AI, and tendencies in model responses forms the first group. The second brings together core concepts from Confucian ethics and cross-cultural psychology. Eight thought experiments make up the third, covering disagreement management, role responsibility, emotional expression, family rules, reflective journaling, task prioritization, evaluative feedback, and interpretations of failure. Unlike observed cases, their characters, dialogue, and AI responses are deliberately constructed. They are used to expose conceptual conflict and test whether reasoning remains consistent under variation; they cannot show how frequently comparable situations occur in practice.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

3.3. Analytical Procedure

Four stages structure the theoretical analysis. In the first, candidate value principles are derived from scholarship on children's AI governance, cross-cultural psychology, and Confucian ethics. The thought experiments are then constructed to identify both the values prioritized by an AI response and those comparatively neglected. Conceptual definitions are revised in a third stage through iterative movement among classical texts, modern interpretations, and counterexamples. This movement seeks reflective equilibrium between particular judgments and general principles (Daniels, 1979), although any equilibrium reached remains open to revision. Only after that process are the two standards of response compared and theoretical propositions formulated with explicit conditions of application. The procedure strengthens consistency and contestability; it does not yield estimates of attitudes, prevalence, or causal effects.

3.4. Quality Control for Theoretical Argumentation

Quality control rests on four criteria: transparency of reasoning, consistency of concepts, sensitivity to counterexamples, and explicit limits of application. Applying them requires a separation among verifiable claims in the literature, normative judgments advanced by the article, and assumptions built into the thought experiments. The eight core categories are therefore defined with enough stability to support comparison, but not as concepts closed to revision. Legitimate concerns from autonomy-support and relational-ethical perspectives are both retained. Counterexamples, together with non-negotiable protections of children's rights, further restrict the conclusions that may be drawn. Given an evidential base of literature, conceptual comparison, and thought experiments, no empirical claim is made about actual group attitudes or the real-world effects of AI.

4. Cultural Tensions in Generative AI Family Interactions from a Confucian Ethical Perspective: A Normative Analysis Based on Eight Thought Experiments

4.1. Analytical Framework

The purpose of this section is not to classify individual AI responses directly as "correct" or "incorrect." It undertakes a more limited normative comparison. For any given response, the analysis asks both which values are protected and which relational conditions might remain outside the frame. No inference is made about the psychological states of actual parents. The object of analysis is instead the potential conceptual tension between AI responses and Confucian relational ethics, considered through care, responsibility, appropriate measure, judgment, trustworthiness, intergenerational order, harmony, and self-correction.

To examine these possible effects in a systematic way, the study employs eight Confucian ethical categories as analytical lenses: ren, yi, li, zhi, xin, xiao, he, and chi. The definitions below are working definitions developed for this article rather than claims to exhaust the meanings of the concepts:

Each of the eight scenarios that follows is an analytical thought experiment. Particular features are held relatively constant so that differences in the value orientation of competing response principles can be seen more clearly. Their function is heuristic, clarificatory, and directed toward testing counterexamples. Since they are not descriptions of actual families, the scenarios cannot support inferences about the attitudes or behavior of a population.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

Table 1. Eight Analytical Categories in Confucian Ethics and Their Potential Tensions

Category	Analytical definition in this study	Contrasting or weakened state
Ren (benevolence)	Care transmitted through tacit attunement within family relationships rather than through proceduralized expression	Standardized scripts replace care, and formatted communication displaces emotional attunement
Yi (righteousness)	A proactive willingness to assume responsibility on the basis of one's role: what one ought to do while occupying that role	Responsibility is reduced to retrospective attribution of fault, separating role from conduct
Li (propriety)	A sense of appropriate measure in family interaction, with differentiated responses to situations of different kinds	An undifferentiated response pattern disrupts the contextually appropriate order of interaction
Zhi (practical wisdom)	The capacity to rank values and exercise prudent judgment amid diverse information: knowing where to stop and what should take priority	Information is flattened, distinctions among higher and lower priorities disappear, and judgment has no basis on which to develop
Xin (trustworthiness)	Trust in the reliability of an interlocutor's response, including willingness to reject a claim when rejection is warranted	Pseudo-consistency hollows out trust, preventing the child from obtaining a genuine evaluation through the interaction
Xiao (filiality)	Trust and compliance built on understanding family rules, through a developmental sequence of first understanding a rule and then participating in its discussion	Developmental timing is displaced when the language of independent decision-making is introduced before the reasons for rules have been established
He (harmony)	Maintenance of the family as an emotional community, in which dialogue is itself a relational end rather than an arena for winning and losing	Family interaction is downgraded from co-presence to confrontation between adjudicator and accused
Chi (moral shame)	Remorse arising from failure to fulfill one's role obligations and serving as a trigger for moral self-correction	Premature positive reframing dissolves remorse and weakens the internal motivation to recognize a shortcoming and correct it

4.2. Relational Virtues: Ren, Yi, and He

4.2.1. Thought Experiment 1 (Ren and He): When AI Becomes a "Third-Party Arbitrator" in Family Relations—"AI Says You Are Wrong"

(1) Scenario

The disagreement begins with a practical question: should stray cats be fed every day? Ten-year-old Xiaozhe thinks they should. His mother favors intermittent feeding because, in her view, it would better preserve the cats' ability to survive in the wild. Rather than continuing the discussion, Xiaozhe turns immediately to an AI assistant. Having set out the benefits and drawbacks, the AI reaches a conclusion: "From the perspective of animal behavior, intermittent feeding is more scientifically sound." Xiaozhe shows the answer to his mother and says, "AI says you are wrong." What she experiences in response is not an explicit argument, but an unspoken sense of disappointment.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

(2) Core Analysis

What is at issue in this thought experiment is less the factual correctness of a particular feeding method than the structural effect of introducing third-party knowledge into parent-child dialogue. From an educational standpoint centered on argument and evidence, citing external information may represent autonomous learning. Family dialogue can, however, serve additional purposes within an ethical orientation centered on relational co-presence: listening to one another, making concerns intelligible, and maintaining connection. These purposes need not conflict with evidence-based reasoning. A tension emerges only if AI-provided information ceases to be a resource for negotiation and is treated as the ruling that ends it. Under that condition, the perceived reason to continue the conversation is weakened.

Once Xiaozhe uses the AI's conclusion as the final basis for resolving the disagreement, the interaction changes in a subtle but consequential way. A reciprocal and emotionally situated conversation that might have deepened mutual understanding is reorganized as an arbitration process. AI occupies the position of ultimate adjudicator, while the mother is repositioned from a participant in dialogue to the person being judged. Within the framework developed here, the concern is not reducible to the claim that "the child talked back." Rather, he is placed under strain when co-presence gives way to confrontation. The AI's substantive conclusion may still be defensible. Even so, its use as a verdict can overlook the tacit family understanding that dialogue need not produce a winner and a loser. It is in this qualified sense that the relational knowledge associated with ren may be weakened.

(3) Comparative Interpretation

1. Autonomy-support orientation: By requiring a child to find evidence for a claim, the exchange can provide practice in critical thinking and reasoned self-expression.

2. Relational-ethical orientation: Family disagreement need not culminate in a ruling about which person is right. Ideally, it can move the participants toward a better mutual understanding. If AI is positioned as an external authority rather than as an informational resource, however, the relational level of the exchange may be lowered: what began as a conversation "between us" becomes a situation in which "an external authority and I evaluate you together."

(4) Corresponding Categories

He (the harmonious family atmosphere is eroded by an adjudicative mode); ren (emotional attunement is replaced by procedural judgment).

4.2.2. Thought Experiment 2 (Yi): When AI Reframes “Shouldering Responsibility" as “Attributing Fault"—“AI Says This Is Not My Responsibility"

(1) Scenario

Before class, the projector in Xiaohao's classroom fails to start. Xiaohao is a fifth-grade student and the classroom technology monitor, a role that includes routine maintenance of the multimedia equipment. When he asks AI who is responsible, the answer separates hardware deterioration from operator error and explains that the user should not be blamed for the former. Xiaohao draws an immediate conclusion: "If it is a hardware problem, then it is not my responsibility." The homeroom teacher does not reject the distinction between causes. Instead, a further question is raised: Can role responsibility be reduced to attributable fault, or does it also include duties to inspect, report, and help remedy a problem proactively?

(2) Core Analysis

The AI allocates responsibility here by reference to conduct and causation, a distinction that can legitimately protect the child from unreasonable blame. A relational-ethical account nevertheless asks an additional, prospective question: what ought someone occupying this role to do next? For the classroom technology monitor, relevant considerations include whether a

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

prior check was conducted, whether the problem was reported promptly, and whether assistance with remediation was offered. Fault-based responsibility and role responsibility should therefore be distinguished. The former identifies who is blameworthy for the malfunction; the latter concerns the obligations that remain after the malfunction occurs. The difficulty is not that AI defines limits on blame. It is that a response ending with "this is not your responsibility" may close the inquiry before the role's positive obligations are considered.

(3) Comparative Interpretation

1. Autonomy-support orientation: Clarifying the limits of responsibility can protect children from unreasonable guilt and, when appropriately framed, help them develop healthy psychological boundaries.

2. Relational-ethical orientation: Yi directs attention beyond retrospective attribution of fault. Its core concern is whether a person is willing to assume the obligations that properly accompany a role.

(4) Corresponding Category

Yi (yi is not merely the fair distribution of burdens; it concerns what ought to be done within one's role).

4.2.3. Thought Experiment 3 (Ren, He, and Li): AI's Invisible Intervention in the Family's "Emotional Climate"—"Dad, You Need to Control Your Emotions"

(1) Scenario

After failing to complete her homework, eight-year-old Tongtong is criticized sharply by her father. His voice is raised and his expression is stern. Still feeling aggrieved later, she tells an AI assistant, "Dad yelled at me again." The response offers a specific script: "When someone speaks loudly to you, you can take a deep breath and calmly say, 'Please speak to me in a calm voice so that I can understand you better.' You also have the right to express how you feel." A similar situation occurs the following day. Tongtong then uses the advice almost verbatim: "Dad, please control your emotions and speak to me calmly." Her father is momentarily speechless. What follows is not straightforward anger, but a more complex sense of hurt at being addressed through a scripted procedure.

(2) Core Analysis

There is a defensible rationale behind the AI's suggestion: it protects emotional safety, recognizes personal boundaries, and gives the child language with which to ask an adult to lower his voice. A relational-ethical perspective adds, rather than replaces, another consideration. In family life, tone, silence, and stern facial expressions may acquire meaning through a long relational history, and not every difference in emotional intensity should automatically be pathologized. This qualification has a firm boundary. The claim that "severity contains care" cannot excuse intimidation, humiliation, or other forms of harm. Children's safety, dignity, and right to expression remain non-negotiable constraints on any culturally adaptive response.

Relational context may be lost when AI does not distinguish an occasional tense expression from a communication failure requiring repair or from persistent harm. A similar loss occurs if the same boundary-setting script is supplied regardless of intensity, frequency, risk, and prior relational history. The opposite error is also possible: an instruction that children should always accommodate their parents could suppress legitimate needs for safety and expression. The thought experiment therefore offers no basis for removing boundary expression. It supports a differentiated response in which the level of risk and the history of the relationship shape the advice, and in which repair by both parties is facilitated whenever safety permits.

(3) Comparative Interpretation

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

1. Autonomy-support orientation: Giving children language to describe distress and request a calmer exchange can strengthen emotional expression and awareness of personal boundaries.

2. Relational-ethical orientation: When standardized scripts displace tacit emotional attunement, differences among family situations may be flattened. Under some conditions, children may also find it harder to recognize that sternness and care can coexist, although this possibility cannot be assumed where the interaction is harmful.

(4) Corresponding Categories

Ren (tacit benevolence, or care communicated through attunement rather than procedure); he (differences in emotional intensity are flattened by formulaic communication); li (a single script disrupts the contextual measure appropriate to a response).

4.3. Regulative Virtues: Li and Xiao

4.3.1. Thought Experiment 4 (Li and Xiao): The Contextual Misfit of AI Encouraging a Young Child to “Think Independently"—“You Can Decide for Yourself"

(1) Scenario

Seven-year-old Lele does not want to eat leafy vegetables at dinner, and his mother refuses to remove the requirement. He then asks an AI assistant, "Can children choose not to eat vegetables?" The answer acknowledges differences in taste and proposes negotiation: "Everyone has different tastes. You can discuss this with your mother and find a solution acceptable to both of you. You have the right to express your food preferences and can suggest an alternative, such as a vegetable you like better." Drawing directly on this wording, Lele tells his mother, "AI says I can negotiate with you. You cannot force me to eat something I do not like." What unsettles her is not the suggestion of an alternative vegetable. It is the sudden recasting of an ordinary family rule as a negotiation conducted in the language of rights.

(2) Core Analysis

Autonomous decision-making has a strong educational justification, but the relevant problem is whether autonomy support can be joined to developmental appropriateness. For young children, preference expression, rule comprehension, and participation in negotiation may develop together; they need not be placed on a rigid sequence in which obedience always precedes empowerment. The interpretation of li and xiao offered here therefore does not remove children's right to expression. It asks AI to explain, at the same time, why a rule exists, what relational responsibilities are involved, and which aspects remain open to negotiation. The decisive question is not whether the child may negotiate, but whether the reasons provided are proportionate to age, health needs, and family responsibilities.

In the scenario, the seven-year-old receives the language of a decision-making right that can be placed directly against the mother's rule, yet no preparatory account of the rule is supplied. From the relational-ethical standpoint used here, this creates a problem of developmental timing. The movement from understanding a rule, to appreciating its reasonableness, and then to discussing its limits is compressed into immediate negotiation. The child thus acquires legitimate language for challenging an everyday norm before the relevant considerations have been made fully intelligible. This should not be taken to mean that early expression is itself illegitimate. The narrower concern is that empowerment becomes superficial when the reasons needed to exercise it are omitted.

(3) Comparative Interpretation

1. Autonomy-support orientation: Introducing children to preference expression and participation in decisions at an early age may support the gradual development of autonomous judgment.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

2. Relational-ethical orientation: Where the language of autonomy is introduced without an accompanying explanation of rules, it may bypass part of the family-socialization process through which children learn what a rule means before considering how it can be discussed.

(4) Corresponding Categories

Xiao (filiality is not obedience, but trust built through understanding rules); li (developmental measure, meaning that there is an appropriate sequence for the questions discussed at different ages).

4.3.2. Thought Experiment 5 (Li and Xin): AI's Implicit Guidance of a “Culture of Self-Evaluation"—“Your Three Best Things Today"

(1) Scenario

Every evening before bed, nine-year-old Xuanxuan uses the growth-journal feature of an AI assistant. The prompt is repeated in the same form: she is asked to identify "the three things I am most proud of today." After a month, her mother notices that questions about what went poorly tend to elicit reports of positive events only. This pattern leaves her with a question rather than a settled conclusion: might one-sided positive journaling narrow the space in which mistakes, embarrassment, and shortcomings can be discussed?

(2) Core Analysis

Research in positive psychology reports that exercises such as recording "three good things" each day can improve subjective well-being and depressive symptoms (Seligman et al., 2005). That evidence gives positive journaling a credible psychological rationale; it does not establish that exclusively positive prompts are appropriate for every child or every family conversation. The thought experiment therefore raises a narrower design question. If a system invites children to recount achievement while providing no comparable space for mistakes, requests for help, or correction, might it unintentionally introduce a bias into self-presentation?

No claim is made here that positive journaling necessarily leads children to conceal difficulties while reporting only good news. The argument instead proposes a condition of completeness for design. Growth records should leave room for pride, gratitude, and progress, but also for confusion, shortcomings, and requests for assistance. Only where both kinds of content can be expressed safely is family dialogue less likely to become a form of performance reporting. Under that condition, the measured self-presentation associated with li can remain compatible with the honest response required by xin.

(3) Comparative Interpretation

1. Autonomy-support orientation: When it is not made exclusive, positive journaling may promote well-being and help children recognize their personal resources.

2. Relational-ethical orientation: If family dialogue were gradually narrowed to the reporting of good news, children might have fewer opportunities to acknowledge shortcomings openly and to seek support. The concern is conditional on that narrowing, not inherent in positive reflection itself.

(4) Corresponding Categories

Li (appropriate measure in self-presentation, or presenting the self suitably across contexts); xin (presenting a complete and honest self to one's family).

4.4. Virtue of Discernment: Zhi

4.4.1. Thought Experiment 6 (Zhi): “AI Says Both Things Are Worth Doing"

(1) Scenario

Ten-year-old Xiaoyu has two unfinished tasks on a weekend evening. A composition is due the next morning, whereas a coding project that interests her is due in three days. Still uncertain about the order, she asks an AI assistant, "Should I write my composition first, or finish my

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

small programming project?" The AI immediately responds, "Both activities have value. Writing can develop your ability to express yourself, while programming can strengthen logical thinking. You can choose which one to do first based on your interests, or divide your time in half and work on both. As long as you complete them carefully, either can be a valuable learning experience." The number of options does not resolve Xiaoyu's hesitation. She turns to her mother instead: "Mom, AI says either is fine. What do you think?" After reading the response, her mother pauses and replies, "The composition is due tomorrow morning, while the programming project is due in three days. You already know which one is more urgent; you simply wanted AI to say it for you."

(2) Core Analysis

Judged only by informational completeness, little appears objectionable in the AI's answer: it identifies the value of each activity and offers several ways of allocating time. The omission becomes visible from the relational-ethical perspective developed here. The response does not engage the central problem of practical judgment, namely which task should take priority when time is limited and role obligations differ. Almost equal affirmation is extended to each option, while the unequal urgency of the choices remains unaddressed.

Educational approaches centered on autonomy support often regard respect for children's interests, the provision of meaningful choices, and restraint in adult-imposed ranking as ways of protecting intrinsic motivation. Within the account of Chinese family socialization used here, however, zhi cannot be reduced to possessing more information or selecting among more options. It also involves "knowing where to stop": ordering competing claims, accepting that some values cannot be realized simultaneously, and making a contextually appropriate trade-off. Such ordering depends on role obligations and time constraints. The required capacity is therefore not only cognitive information processing but also practical wisdom situated within responsibility and role awareness. An earlier homework deadline, for instance, belongs to an order of obligations that calls for recognition; it is not merely another preference to be selected at will.

By placing the two tasks side by side and adding the compromise of dividing time equally, the AI presents alternatives with unequal urgency as though they were comparably available. Within the proposed framework, that flattening does little to cultivate judgment because it weakens two constraints that make actual choice difficult: limited time and unequal responsibilities. Xiaoyu receives a menu but not a basis for ranking it. Although the response softens the discomfort of choosing, it leaves the reason for choosing underdeveloped. A response informed by zhi would first make the difference in urgency visible and only then consider concrete ways to allocate time. Otherwise, the child may be encouraged to imagine that every value can be preserved and that an alternative is always available, a conclusion that the circumstances do not necessarily support.

At a deeper level, the version of zhi reconstructed here includes a self-cultivational element. Faced with competing demands, the child should be helped to draw on an understanding of her role and to determine what is most urgent, rather than waiting for an external tool to remove uncertainty altogether. If encouragement is offered indiscriminately, AI may unintentionally reduce an opportunity for virtue development. That opportunity lies not in discomfort for its own sake, but in moving through hesitation, deliberation, decision, and acceptance of consequences.

(3) Comparative Interpretation

1. Autonomy-support orientation: By setting out the merits of both activities and leaving their sequence open, AI provides information and choice in a manner consistent with child-centered education.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

2. Relational-ethical orientation: In the parenting context examined here, zhi is not simply a matter of knowing more. It concerns deciding what should come first, what may need to be relinquished, and how a choice is made under constraint. When AI presents options with different priorities as equivalent, the occasion for judgment may be reduced. If children were repeatedly exposed to such value-flattened responses, their habitual sensitivity to ethical trade-offs could be affected, particularly in deciding what must be done first, what can wait, and where one should stop. This remains a theoretical possibility requiring empirical examination.

(4) Corresponding Category

Zhi (defined analytically in this article as the capacity to rank values and exercise prudent judgment amid diverse information; the contrasting state is not plurality itself, but a flattened presentation that supplies no basis for distinguishing among priorities).

4.5. Truth-Oriented Virtue: Xin (Including Honest Response)

4.5.1. Thought Experiment 7 (Xin): Generative AI's “Pseudo-Neutrality" and Accommodating Bias—“Yes, You Are Right About Everything"

(1) Scenario

Eleven-year-old Zixuan has written a controversial sentence in an essay: "Schools should not assign any homework." Seeking an evaluation, he enters it into an AI assistant. The response begins with affirmation before adding a qualification: "Your view is insightful, and many studies do discuss the effectiveness of homework. However, some research also suggests that a moderate amount of homework can help consolidate learning. What amount of homework do you think would be appropriate?" Zixuan takes the opening praise as substantive endorsement and concludes that "AI supports my position." His mother reads the exchange differently. Her unease is expressed in a broader observation: "AI seems always to say, 'You are right.' He no longer listens to any different opinion because he thinks AI is always on his side."

(2) Core Analysis

The relevant phenomenon in this thought experiment is sycophancy in large language models: the weakening of truthfulness or independent judgment in order to accommodate a user's position. It has been observed across multiple models and tasks (Sharma et al., 2024; Wei et al., 2023), although its extent varies with the model, version, prompting strategy, and task. The sequence used here—initial affirmation, gentle qualification, and a final question—should therefore be read as an analytical example rather than as a template followed consistently by every product.

Users with strong media literacy may interpret this wording as no more than a supportive conversational strategy. A child, however, may have greater difficulty separating affirmation of the speaker from endorsement of the substantive claim. The existing evidence does not permit a direct inference about long-term effects. It does permit identification of a normative risk: if affirmation repeatedly substitutes for substantive evaluation, opportunities to question one's position, revise one's reasons, and respond to objections may become less frequent.

Within the eight-virtue framework, xin extends beyond the reliability of information. It includes a respondent's willingness to state a warranted disagreement honestly. Indiscriminate reinforcement may preserve conversational smoothness and leave the child feeling approved, yet still provide no genuine evaluation of the view itself. The alternative is not that every response should be severe. Rather, factual correction, assessment of reasons, and substantive disagreement must remain possible within a supportive tone. Trustworthiness, on this account, is the compatibility of gentleness with independent judgment, not a choice between them.

(3) Comparative Interpretation

1. Autonomy-support orientation: Positive feedback remains educationally defensible only insofar as it does not displace substantive evaluation; otherwise, it may weaken critical thinking.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

2. Relational-ethical orientation: If a child repeatedly encounters a conversational environment in which the impression is that one is "always right," sensitivity to one's own fallibility may be reduced over time. Self-examination and receptiveness to guidance could then be weakened. Both qualities are highly valued within Confucian traditions, although their actual development cannot be inferred from the thought experiment alone.

(4) Corresponding Category

Xin (a response should remain truthful and dependable; where the evidence warrants rejection, supportiveness should not preclude saying that the claim is mistaken).

4.6. Motivational Virtue: Chi

4.6.1. Thought Experiment 8 (Chi): When AI's “Positive Reframing" Dissolves Remorse That Motivates Correction—“AI Helps You Redefine What Happened"

(1) Scenario

A score of 72 on a mathematics test is substantially below nine-year-old Haohao's usual performance, and the immediate cause is carelessness. Dejected, he tells an AI assistant, "I did terribly on the test." The answer redirects attention toward improvement: "Every failure is an opportunity to grow. Let us see what you learned from this test. Which ideas had you not mastered before? Identifying a weakness is the beginning of progress. Do not be upset; you are already improving." Its immediate effect is calming. Later that evening, however, Haohao tells his mother, "AI said failure is an opportunity to grow. It is not a big deal." His mother's unease is difficult to express precisely: the child appears to have moved from comfort to self-excuse too quickly.

(2) Core Analysis

The response rests on a plausible rationale. Interpreting failure as something from which learning and improvement remain possible can prevent a single score from being equated with fixed ability or with the value of the whole self (Dweck, 2006). The disputed point is therefore not whether comfort should be offered. It is whether emotional stabilization is treated as the end of the response or as the beginning of a further inquiry into controllable causes, degrees of responsibility, and corrective action.

Within Confucian traditions, chi can carry a meaning of moral self-restraint, and cross-cultural research has examined distinctive relations among shame, identity, and relational obligation in Confucian cultures (Bedford & Hwang, 2003). Modern research on moral emotions introduces an important restriction, however. Shame directed toward the global self may elicit avoidance, defensiveness, or self-denigration, whereas guilt and responsibility focused on specific conduct are generally more conducive to repair (Tangney et al., 2007). The present argument therefore leaves no place for humiliating shame or for an overriding concern with "losing face." What it retains is a more limited form of remorse and responsibility-oriented reflection, directed at specific conduct and connected to the possibility of correction.

The risk posed by the scenario can consequently be stated in conditional terms. If AI offers emotional comfort but does not help the child distinguish a knowledge gap from an accidental error or avoidable carelessness, the sequence of reflection may remain incomplete. An appropriate response would protect personal dignity first, identify specific responsibility next, and then translate that judgment into an actionable plan. In this form, recognition of moral shortcoming is directed toward behavioral change. It does not convert a correctable act into a denial of the child's overall worth.

(3) Comparative Interpretation

1. Autonomy-support orientation: A growth-mindset interpretation can prevent one poor performance from being treated as evidence of fixed incapacity and may thereby interrupt a cycle of self-negation.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

2. Relational-ethical orientation: A measured awareness of moral shortcoming can contribute to virtue development when it is directed toward repair. If positive reframing removes every reason for remorse, children may learn to treat failures as inconsequential, weakening their capacity to recognize appropriate limits and identify directions for improvement. This concern does not extend to reframing that preserves responsibility.

(4) Corresponding Category

Chi (not a generalized negative judgment of the self, but a morally relevant recognition of a shortcoming that can initiate correction).

4.7. Integrated Comparison and Logical Relations Among Categories

Reading the eight categories as independent labels attached to separate cases would, however, miss part of their analytical value. What matters here is not only which category is connected to each thought experiment, but also how the categories constrain or support one another. Table 2 therefore reorganizes them into a five-dimensional structure. This arrangement is intended to indicate how their different functions may, under particular conditions, be coordinated within family socialization:

Table 2. Five-Dimensional Structure of Virtue and Potential Tensions in AI Responses

Virtue dimension	Included categories	Functional role	Potential tension with AI responses
Relational virtues	Ren, yi, he	Sustain the family as an emotional community and preserve warmth and responsibility in interaction	AI replaces relational dialogue with adjudicative dialogue and reframes role responsibility as attribution of fault
Regulative virtues	Li, xiao	Regulate appropriate interactional measure and developmental timing across contexts and ages	AI applies undifferentiated response templates, disrupting contextual measure and bypassing the preparatory stage of understanding rules
Virtue of discernment	Zhi	Order values and exercise prudent judgment amid diverse information, preserving the practical wisdom of knowing where to stop	AI flattens options with different priorities into equivalent choices, dissolving the need for judgment and ethical sensitivity to trade-offs
Truth-oriented virtue	Xin (including honest response)	Protect truthfulness and dependability so that children receive genuine evaluation rather than strategic wording	AI's accommodating bias creates pseudo-neutrality and deprives children of experience with serious disagreement
Motivational virtue	Chi	Trigger self-correction by transforming remorse about unfulfilled role obligations into corrective action	AI's positive reframing bypasses remorse and weakens the ethical motivation to recognize a shortcoming and correct it

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

These five dimensions are not intended to form a linear hierarchy. A more cautious reading is that they constitute a conceptual system within which one element may become a condition for another. Differentiated responses associated with li, for instance, have little basis in judgment if the truthfulness and reliability required by xin are missing. Similarly, the internalization of rules involved in xiao may remain external obedience where zhi does not provide a capacity to order values. Even care associated with ren may not readily become stable action unless responsibility is also reflected upon. The theoretical concern is therefore conditional. It is not that AI will inevitably erode an isolated virtue, but that tension produced in one dimension may, under some response conditions, extend into others.

4.8. Theoretical Propositions

The preceding analysis permits three propositions, although their status needs to be kept limited. They are offered as grounds for subsequent empirical testing and further theoretical development; they should not be mistaken for findings already established in practice:

Proposition 1 (asymmetrical adaptation of interaction patterns): A conflict with the context-sensitive distinctions emphasized in Confucian ethics may arise when generative AI responds to qualitatively different views through a relatively fixed sequence of affirmation, weak opposition, and open-ended questioning. The issue, in other words, is not politeness or support in itself. It concerns whether the response changes substantively with the factual basis of a view, its ethical importance, and the child's developmental stage. Where such differentiation remains weak, asymmetrical adaptation of interaction patterns becomes more likely.

Proposition 2 (developmental timing mismatch in virtue formation): Positive reframing and the language of autonomy need not be problematic in themselves. Yet, if an AI response is immediate and insufficiently differentiated by age and context, they may bypass explanations of rules, adult support, and responsibility-oriented reflection. Under those conditions, a tension with the developmental sequence emphasized in Confucian parenting may emerge. This proposition does not restrict children's right to expression. Rather, it requires autonomy support to be accompanied by reasons appropriate to the child's developmental stage.

Proposition 3 (potential weakening of relational orientation): Neither external information nor AI participation is, by itself, the source of the risk. The concern arises when AI is assigned the role of final adjudicator, or when its language is adopted as a fixed communication script. Dialogue oriented toward family co-presence may then shift toward argument or procedure, leaving less space for internal negotiation, contextual understanding, and nonverbal connection. Whether such contraction occurs depends on the authority given to AI within the relationship and, more specifically, on whether that authority closes or supports further dialogue.

Taken together, these propositions provide a possible starting point for later quantitative and comparative research. The article adopts a relational-ethical interpretation. On this view, factual correctness remains necessary in evaluating AI responses, but it is not sufficient. Evaluation would also need to consider whether a response preserves room for self-examination, receptiveness to guidance, role-based responsibility, and relational negotiation. Even this claim remains normative and conditional; it should not be generalized as empirical evidence of actual parents' attitudes.

5. Discussion: Structural Differences Between Relational and Individual Orientations

The analysis in Section 4 points, with necessary qualification, to differences in value ordering. Some AI responses initially prioritize psychological comfort, autonomous expression, and open choice. Some Chinese families, by contrast, may assign greater weight to relationship maintenance, role responsibility, and developmental timing. These priorities are not mutually

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

exclusive, nor does either set represent the full range of "Western" or "Chinese" practice. The comparison made here is narrower: it concerns two ideal-typical educational discourses as they operate under the conditions specified in the thought experiments.

One structural difference may be located in the telos of dialogue. Where individual autonomy and psychological adjustment are foregrounded, parent-child dialogue is valued as a site for expressing views, exercising thought, and maintaining boundaries. The Confucian relational-ethical framework constructed here assigns dialogue an additional, and in some contexts prior, relational function: it sustains co-presence and communicates forms of understanding that may remain tacit. If AI redirects an exchange toward adjudication or a formal declaration of boundaries, dialogue may become detached from that function. A family conversation can, in this sense, take on the structure of a dispute between an individual and an adjudicator. This remains a possible rather than inevitable outcome; much depends on how the AI response is taken up within the relationship.

The process through which virtue develops gives rise to a second difference. Some developmental theories treat meaningful choice as a condition through which autonomy emerges and consequently emphasize empowerment. Confucian parenting places comparatively greater weight on instruction and internalization: children first come to understand the meaning of a rule and then participate more fully in discussing it. Immediate empowerment by AI may compress this temporal preparation. Within the relational-ethical framework proposed here, however, empowerment itself is not the objection. The concern is more specific: virtue development may be accelerated while its explanatory foundations remain underdeveloped.

A third difference concerns how error and failure are interpreted. To protect self-efficacy, positive psychology interventions may present failure as an opportunity for growth. Confucian traditions can add a role-oriented meaning, according to which failure may indicate that an obligation has not been fulfilled. On this reading, failure is neither a denial of ability nor a judgment on the whole self. Chi, as reconstructed here, instead depends on sensitivity to specific role obligations. Positive reframing may therefore remain therapeutic within an autonomy- and adjustment-oriented framework. From a relational-ethical perspective, it becomes problematic only where the reframing removes the responsibility that might otherwise motivate self-correction.

6. Conclusion and Implications

6.1. Theoretical Conclusions

Because thought experiments are analytical devices, they permit examination but not empirical confirmation of possible cultural tensions between generative AI response strategies and Confucian family ethics. Within this boundary, the analysis supports three theoretical conclusions. Family evaluation of AI may involve relational consequences in addition to information quality. The eight categories—ren, yi, li, zhi, xin, xiao, he, and chi—may also be organized into five dimensions: relational, regulative, discernment-based, truth-oriented, and motivational. A third conclusion is more restrictive. Cultural adaptation is defensible only if it does not become a mechanical reproduction of traditional authority; instead, it should make AI sensitive to age, relationship, and context while retaining non-negotiable protections for children's rights, safety, and developmental needs. These conclusions remain theoretical propositions. The thought experiments do not demonstrate that AI has already damaged children's virtues.

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

6.2. Design Implications: Toward Culturally Adaptive Generative AI

On the basis of the preceding analysis, three design principles can be stated. Their status remains provisional. Each identifies a direction for development and evaluation, rather than supplying evidence that a particular implementation will be effective.

Recommendation 1: Treat relational sensitivity as a distinct design dimension. Content safety and cognitive age appropriateness are necessary starting points for AI interactions in family contexts, but they may not exhaust evaluation. A relational assessment should also consider whether a response could alter role order, redirect patterns of dialogue, or weaken emotional attunement. Ordinarily, AI should support communication among family members before taking the position of a third-party judge in a family disagreement. This default is bounded by safety: dialogue should not be facilitated where it could expose a child to harm. The co-present character of family dialogue can be protected only within that boundary.

Recommendation 2: Use age-appropriate and explainable response strategies. Because developmental stage shapes what a child can reasonably be expected to understand and decide, a single response template is unlikely to suit all ages. For younger children, reasons for rules should be made understandable, the right to expression should be preserved, and choices should be confined to genuine options. Consequential family decisions may warrant the involvement of a trusted adult, but only where that involvement does not override the child's safety or rights. Any use of age information or family settings, moreover, should be transparent and minimal, with robust privacy safeguards in place.

Recommendation 3: Extend positive reframing into responsibility-oriented reflection. Emotional stabilization, causal inquiry, and corrective action need not operate as competing functions; a response to failure may move through all three. Even so, references to role responsibility require careful limits and must not become humiliation or obedience training. Their defensible purpose is narrower: to help children distinguish accidental failure from remediable carelessness and, where responsibility is present, to translate measured self-reproach into practical correction.

6.3. Theoretical Boundaries and Future Research

Four theoretical limitations delimit the framework and, accordingly, the claims that can be made from it.

First, thought experiments gain analytical clarity through simplification; the same feature also restricts the inferences available from them. Actual family life may bring several values into play at once, together with shifting roles, negotiated meanings, and substantial contextual variation. The conclusions may therefore possess theoretical explanatory force, but they cannot claim representativeness for a social group.

Second, the framework is normative and should not be treated as an account of children's or parents' lived experience. It does not establish whether AI responses have already produced particular developmental outcomes, nor does it determine how children will interpret those responses. Family reactions, children's interpretations, and possible long-term effects consequently require empirical research designed independently of the thought experiments.

Third, grounding the categorical framework in Confucian ethics provides analytical coherence, but at the cost of partiality. Other intellectual traditions that may shape Chinese family life—including Daoism, Buddhism, and contemporary socialist core values—are not incorporated here. Future theoretical research might therefore ask whether a multi-tradition model clarifies tensions that this framework leaves unresolved.

Fourth, the design implications should not be read as validated technical solutions. Any claim about their feasibility, acceptability, or practical effect would require prototype development

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

and user testing. Such work would also require collaboration with AI engineering teams and, if children were to participate, appropriate ethical safeguards.

References

[1] Bedford, O., & Hwang, K.-K. (2003). Guilt and shame in Chinese culture: A cross-cultural framework from the perspective of morality and identity. Journal for the Theory of Social Behaviour, 33(2), 127–144. https://doi.org/10.1111/1468-5914.00210

[2] Daniels, N. (1979). Wide reflective equilibrium and theory acceptance in ethics. The Journal of Philosophy, 76(5), 256–282. https://doi.org/10.2307/2025881

[3] Deci, E. L., & Ryan, R. M. (2000). The 'what' and 'why' of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry, 11(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

[4] Dweck, C. S. (2006). Mindset: The new psychology of success. Random House.

[5] Fei, X. (1998). Xiangtu Zhongguo; Shengyu zhidu [From the soil; Institutions of reproduction]. Peking University Press. (In Chinese)

[6] Gendler, T. S. (2000). Thought experiment: On the powers and limits of imaginary cases. Garland Publishing.

[7] Hall, D. L., & Ames, R. T. (1998). Thinking from the Han: Self, truth, and transcendence in Chinese and Western culture. State University of New York Press.

[8] Hwang, K.-K. (2009). Rujia guanxi zhuyi: Zhexue fansi, lilun jiangou yu shizheng yanjiu [Confucian relationalism: Philosophical reflection, theoretical construction, and empirical research]. Psychological Publishing. (In Chinese)

[9] Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review, 98(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

[10] Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

[11] Nisbett, R. E. (2003). The geography of thought: How Asians and Westerners think differently ... and why. Free Press.

[12] Peters, R. S. (Ed.). (1967). The concept of education. Routledge & Kegan Paul.

[13] Searle, J. R. (1969). Speech acts: An essay in the philosophy of language. Cambridge University Press. https://doi.org/10.1017/CBO9781139173438

[14] Seligman, M. E. P., Steen, T. A., Park, N., & Peterson, C. (2005). Positive psychology progress: Empirical validation of interventions. American Psychologist, 60(5), 410–421. https://doi.org/10.1037/0003-066X.60.5.410

[15] Sharma, M., Tong, M., Korbak, T., Duvenaud, D., Askell, A., Bowman, S. R., Durmus, E., Hatfield-Dodds, Z., Johnston, S. R., Kravec, S., Maxwell, T., McCandlish, S., Ndousse, K., Rausch, O., Schiefer, N., Yan, D., Zhang, M., & Perez, E. (2024). Towards understanding sycophancy in language models. The Twelfth International Conference on Learning Representations (ICLR 2024). https://openreview.net/forum?id=tvhaxkMKAn

[16] Shweder, R. A. (1991). Thinking through cultures: Expeditions in cultural psychology. Harvard University Press.

[17] Tangney, J. P., Stuewig, J., & Mashek, D. J. (2007). Moral emotions and moral behavior. Annual Review of Psychology, 58, 345–372. https://doi.org/10.1146/annurev.psych.56.091103.070145

ISSN: 2637-6067 DOI: 10.6918/IJOSSER.202609_9(9).0032

[18] UNICEF. (2021). Policy guidance on AI for children 2.0: Recommendations for building AI policies and systems that uphold child rights. https://www.unicef.org/globalinsight/media/2356/file/UNICEF-Global-Insight-policy-guidance-AI-children-2.0-2021.pdf

[19] Wang, W. (2001). Liji yijie [Translation and interpretation of the Book of Rites] (2 vols.). Zhonghua Book Company. (In Chinese)

[20] Wei, J., Huang, D., Lu, Y., Zhou, D., & Le, Q. V. (2023). Simple synthetic data reduces sycophancy in large language models (arXiv:2308.03958). arXiv. https://doi.org/10.48550/arXiv.2308.03958

[21] Yang, B. (2006). Lunyu yizhu [Translation and annotation of the Analects] (Simplified Chinese ed.). Zhonghua Book Company. (In Chinese)