

Cyclic reinforcement and the social superego: a hypothesis on how family and culture sustain harm, order, nurture and defection across generations

Paper IV of the Peace Papers, six papers on one framework; each stands alone. Website version, working draft v11, 30 October 2026. Devon Summerhayes, London.

Abstract

Harm, order, nurture and defection repeat across generations and through peer groups. This article proposes cyclic reinforcement, a model in which a loop's output strengthens its own inputs in the next generation, in four forms: cyclic harm, cyclic order, cyclic nurture and cyclic defection. In each, two inputs shape the child: parenting, and the social superego, the internalised average of how the people around a child behave, weighted by their influence; the adult then parents and adds to the norm. Family and wider setting are coupled loops that buffer each other, so that a nurturing setting can take up a child whose family input is weak. There are two routes into harm: order whose norm licenses a target, and defection, a norm kept for others and broken for oneself, hidden behind a public face or openly declared, which gains influence when it goes unsanctioned. The model is positioned in family theory and illustrated by three measured cases: the culture of honour, cultural tightness and authoritarianism, and defection in public and in private. Genetic variation may modify sensitivity to the environment; MAOA, the worked example, is the weakest evidence and is treated as such. The model names three exits from the loop, another adult or group, reflective capacity, and forgiveness, at least ceasing to retaliate, and draws practices from them. It states seven hypotheses, each with a design and with what would count against it, and labels every descriptive claim Experienced, Measured or Interpretive. Its contribution is integration and a framing, not a new mechanism; whether it is right depends on tests that others can run.

In the series, the paper supplies the social superego and the carriers that pass norms on, which paper VI extends to trained machines. It takes the felt and the understood perspectives from paper I and forgiveness from paper II, and its revenge form is the vendetta that paper III sets beside the revenge liar. Paper V is the protocol for a systematic review designed to test its worked example, MAOA and childhood adversity.

1. Introduction

Why do cruelty, neglect and conformity to harmful norms repeat across generations, and why do some people escape the pattern while others do not? And why does the same strictness produce, in one family or society, an adult who upholds order, and in another one who does harm? And why do some people hold others to a norm from which they exempt themselves, some behind a show of keeping it and some openly? Research offers part of the answer: maltreatment raises later antisocial risk (Widom, 1989), attachment patterns pass from parent to child (van IJzendoorn, 1995), some genetic variants appear to modify that risk (Caspi et al., 2002), what people see others do shapes what they do (Cialdini et al., 1990), and people will keep up an appearance of fairness while serving themselves (Batson et al., 1997). This article ties those parts into a single picture and states where each rests on evidence. It is offered as a contribution to family theory: a model of how the family and the wider setting act together, as two coupled loops, on what each generation passes to the next, whether harm, order, nurture or defection.

1.1 Approach (Experienced and Interpretive)

Experienced. The model comes from my observation of how influence passes between generations, in families and in wider groups, together with self-examination, and was afterwards checked against published research. It is observational and interpretive. It does not draw on clinical practice or on

case material. To keep evidence and opinion apart, every descriptive claim in this article belongs to one of three modes, Experienced, Measured or Interpretive, defined in Table 1 (the labelling discipline of paper I, section 3); section headings state the modes they use, except the Introduction and Conclusion, whose headings the journal fixes, and paragraphs that mix modes label each.

Table 1. Three modes of claim used in this article.

Mode	Meaning	Standard it must meet
Experienced	First-person report: my own observation and self-examination	Honesty and consistency with other accounts: testimony, not proof, and never used as evidence for a Measured claim
Measured	Findings that were measured and can be checked	Evidence, replication and effect size, from a cited source with its limits stated
Interpretive	A way of reading the measured facts, or a model built on them	Coherence and usefulness, whether it clarifies more than it distorts: argued openly, offered as a hypothesis, open to being wrong

Generative AI, Claude (Anthropic, 2026), was used, Sonnet 5.5 for the first draft, Haiku 5.5 and Fable 5.1 for revision and review, and Opus 5.5 for revision, including this version, to help draft and edit the text, to build the figures, to find and summarise literature, with every source then checked against its publication record, to check references, citations and internal consistency, and to convert the references to APA style. Some connections in the text, the measured cases for the order and defection loops, the overvaluation input of the defection loop, and four of the hypotheses in their present form (H1, H4, H6 and H7), were suggested by AI in discussion and adopted by me; the defection loop itself, its name, its two faces and the defecting parent as its input are mine. I directed the work, reviewed all content and take full responsibility for it.

2. The model: cyclic reinforcement in four forms (Interpretive)

The model borrows the ego-state vocabulary of transactional analysis (Berne, 1961, 1964): Parent, Adult and Child states, with the Child divided into a natural Child (later called the Free Child, the name used here) and an Adapted Child that has learned to comply. It adds one construct, the social superego, defined here as the child's internalised average of how the people who influence it behave, weighted by their influence, which becomes what the child expects of itself and others and punishes itself for breaking. Its content, called norm content below, is what that average conduct permits and expects, including whom it licenses as a target. The average is of behaviour, the descriptive norm of Cialdini et al. (1990); internalisation is what turns it into an expectation with a felt penalty: in this model the descriptive norm becomes an injunctive one held as one's own. Cialdini et al. keep the two apart, and H1 measures them apart; H1 does not measure the weighting by influence, which belongs to the definition and not yet to any test. It follows Freud (1930/1961a), who extended his own superego (Freud, 1923/1961b) from the family to a cultural superego, but it defines the group-level structure as an in-group average and places it in a loop. The model is called cyclic reinforcement because the loop's output, the adult, strengthens its own inputs in the next generation, the parenting and the norm; reinforcement here means that self-strengthening, not the operant sense used in coercion theory. The loop comes in four forms with one structure (Figure 1). Three differ in what the Parent does, what the norm expects, and what the adult passes on; the fourth, defection, differs in the child's relation to the norm, which the child learns, by example or by exemption, does not bind it.

Cyclic harm (Figure 1A). A Free Child receives coercive parenting, harsh, arbitrary and inconsistent, and absorbs a norm that permits harm. The child becomes an Adapted Child, then an adult. As an adult the person parents the next generation in the same manner (vertical transmission) and adds to the norm that surrounds peers (horizontal transmission). The pattern resembles, as an analogy only, what Dawkins (1982) called an extended phenotype, an organism's effect on its surroundings; his concept is of genes and the pattern here is learned, so the borrowing is of the shape, not the mechanism. Measured. The route from compliance to harm is the one coercion theory and social learning describe: a child who is coerced learns that coercion works and is normal (Bandura, 1977; Patterson, 1982). Interpretive. The child carries it into its own parenting and its

dealings with others. How one person's conduct moves an in-group average is assumed to work through the same observation and reward, repeated across many members; the model does not test that step.

Cyclic order (Figure 1B). The same structure with a controlling but consistent, rule-based Parent and a norm of order. The child becomes an Adapted Child who is pro-order: obedient, conventional, and as an adult parents by rules and enforces the norm. Order reproduces itself, with its benefits, predictability, safety and low crime, and its costs, low autonomy and intolerance of deviance. This loop answers a question a model with only harm and nurture cannot: why compliance, which is the Adapted Child's defining feature, is not itself antisocial.

Cyclic nurture (Figure 1C). The same structure with a nurturing Parent and nurturing norms: the Free Child can develop freely, becomes a self-directed adult, and passes on nurture. These three loops differ in content, not in mechanism, so the model predicts that changing either input (parenting or norm) changes what is passed on.

Cyclic defection (Figure 1D). The same structure with a Parent who defects, or one who holds the child above others, or both. A defecting Parent teaches by example: the child sees the rule kept for others and broken by the parent, behind a show of keeping it or openly, and learns the behaviour, and often the mask with it, through the observation and reward that social learning describes (Bandura, 1977); this route needs no warmth. An overvaluing Parent teaches by exemption: warmth without limits, with the child held above others, the indulgent pattern of Maccoby and Martin (1983) with the overvaluation that Brummelman et al. (2015) measured; this route can run even where the parent keeps the rule. The model predicts that the two compound, with transmission strongest where both are present. Either way the child learns the norm and learns that it is for others. It neither complies, as the Adapted Child does, nor develops within limits given with warmth, as the nurtured child does. As an adult the person keeps others to the norm and exempts themselves. This is defection in the sense of the cooperation literature (Nowak & Sigmund, 1992), serving oneself while others keep to the rule, which pays only because they do; the person's relation to the norm, held for others and broken for oneself, is called self-exemption here, and the person the exempted self. The retaliation of the culture of honour, below, which game theory also counts as defection, is a different thing: it is licensed by its norm and belongs to the revenge form of cyclic harm. Defection has one of two faces: hidden, the norm kept in public and broken in private, or open, the norm declared not to apply to oneself, or to be for the weak. The model predicts that sanction, a penalty for breaking the norm, formal or informal, decides the face: hidden where breaking the norm is punished, open where it is not, or where the person has power enough to escape it. The adult passes defection on vertically, by the same two routes: defecting where their own child can see, which teaches the rule and the mask together, or holding the child above others; and horizontally, since the social superego is weighted by influence, and an exempted adult who gains influence counts for more in the average the next generation internalises. Measured. In 565 children aged 7 to 12, followed over four waves six months apart, parental overvaluation predicted narcissism and a lack of parental warmth did not, while warmth predicted self-esteem (Brummelman et al., 2015). Among 186 children aged 3 to 7, those aged 5 to 7 who had just been lied to by an adult were more likely to peek at a toy they had been told not to look at, nearly 80 against about 60 per cent, and more likely to lie about it afterwards, nearly 90 against about 60 per cent of those who peeked; the youngest were not affected (Hays & Carver, 2014). Among 379 young adults, those who recalled more lying by their parents in childhood reported lying to their parents more as adults (Setoh et al., 2020), a retrospective and correlational finding. Interpretive. Narcissism is a measured neighbour of the exempted self, not the same thing: self-exemption is a relation to a norm, and the model's claim is that it is passed on. The measured cases for the two faces and for the weighting are given in their own section below.

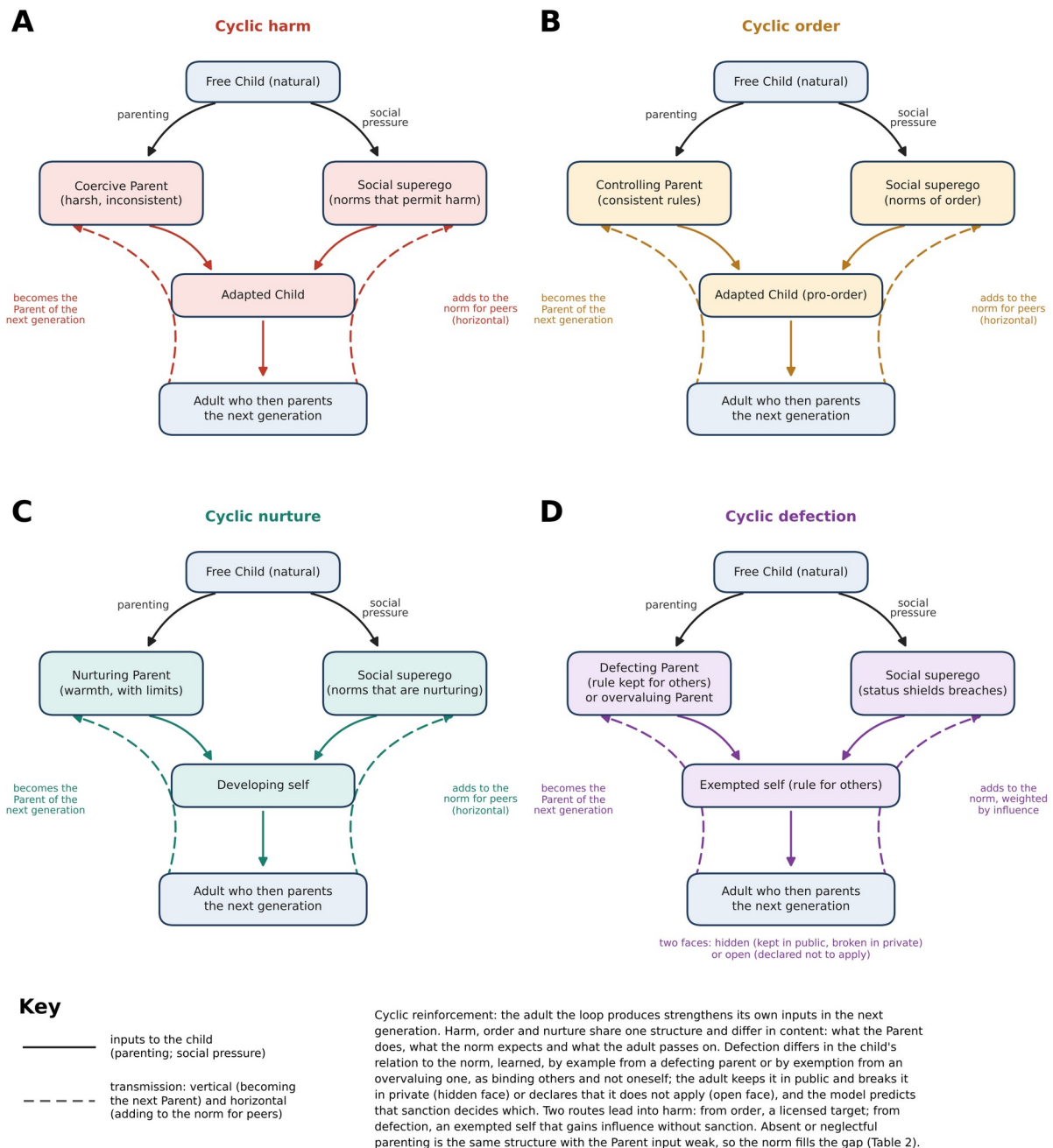


Figure 1. Cyclic reinforcement in its four forms: cyclic harm, cyclic order, cyclic nurture and cyclic defection. Panel A: cyclic harm. Coercive parenting and a norm that permits harm shape an Adapted Child, who parents the next generation and sets the norm for peers. Panel B: cyclic order. Consistent, rule-based control and a norm of order shape a pro-order Adapted Child, who passes on the rules. Panel C: cyclic nurture. Nurturing parenting and nurturing norms let the Free Child develop, who then parents and models nurture. Panel D: cyclic defection. A defecting or overvaluing parent and a setting in which status shields breaches teach the norm as one for others, by example or by exemption; the adult keeps others to it and exempts themselves, with a hidden face (the norm kept in public, broken in private) or an open one (the norm declared not to apply), and adds to the norm with the weight of their influence. Solid arrows show the inputs to the child; dashed arrows show vertical transmission (becoming the parent of the next generation) and horizontal transmission (adding to the norm for peers). The key states the two routes into harm and how absent or neglectful parenting fits.

What separates order from harm. Interpretive, with Measured support. The split proposed is not the amount of control but its kind and its content. On the parenting side the literature has two constructs the split draws on: Baumrind's (1991) authoritarian pattern, high demands with low responsiveness, which values obedience, and Patterson's (1982) coercive process, harsh and

escalating, in which inconsistency is a feature of the exchanges he observed. Consistency as the dividing property is the model's addition; neither source measured it as such, and H1 treats it as a claim to test. On the norm side the split is norm content: whether the norm licenses harm to anyone (retaliate, take) or not (obey). Measured. Duckitt's (2001) dual-process model, built from cross-sectional survey paths, gives the same split a measured shape: punitive socialisation is linked, through social conformity and a view of the world as dangerous, to right-wing authoritarianism, and unaffectionate socialisation, through tough-mindedness and a view of the world as a competitive jungle, to social dominance orientation; the paths are associations, not a demonstrated sequence. Interpretive. The first is the order loop's adult; the second runs toward the profile described in the section on Dark Tetrad traits. The two loops are neighbours with one switch between them: an order that carries a licensed target. Altemeyer's (1981) authoritarianism has three parts, submission to authority, conventionalism, and aggression in the name of authority, and the third is the switch. Where the order licenses harm to a target, the harm is felt as duty rather than guilt, and cyclic order becomes cyclic harm.

A second route into harm. The switch is one route into harm; defection is a second. An exempted self feels no penalty at its own breach, so the harm it does is felt as advantage rather than as duty or guilt. While sanctions hold, the face stays hidden and the harm private. But an exempted adult who gains influence without sanction counts for more in the in-group average, and where such breaches become common and nothing follows, what one generation did while exempting itself the next does in obedience to the norm: cyclic defection becomes cyclic harm. The first route turns on a licensed target, the second on an exempted self. H7 (Table 3) tests the second, and with it the weighting by influence that the definition of the social superego assumes.

What the model does not say. It does not say that every person who is parented harshly becomes harmful; most do not. It does not say the loop is fixed, or that people are only products of it. Nurturing includes firm limits given with warmth (Baumrind's authoritative pattern); the order loop is control without warmth (her authoritarian pattern); the harm loop is control that is harsh and inconsistent (Patterson's coercive process); the defection loop is a parent who defects, or warmth without limits with the child held above others (the indulgent pattern), or both; and absent or neglectful parenting is the low-gain case described below (gain here is the strength with which a loop turns what it receives into what is passed on). Free Child, Adapted Child and exempted self are descriptive labels, not measured entities, and the model is not offered as a clinical tool. The loop describes a tendency, not a track. Whether a pattern counts as adapted depends on whose view is taken: the child, the parent and an outside observer may each describe the same behaviour differently, and the labels here are the observer's, not the child's. Exits exist, such as repair after a rupture, a steady relationship with another adult, and peers or schools with nurturing norms (Table 2), and the capacity to see the loop from within it (described below).

Who the loop does not capture. Not everyone is captured, and the model separates two cases that look alike from outside, after setting aside a third that is often confused with them. The rebel is the third: reacting against a norm is still being shaped by it, the rebellious face of the Adapted Child in later transactional analysis, and the loop runs inverted with the inversion as its content. The child who does not register the input is the low-susceptibility child of differential-susceptibility theory (Belsky & Pluess, 2009), and the model's prediction for this child cuts both ways: transmission in both directions is weakest, the least harm from cyclic harm and the least help from cyclic nurture, the mirror of Table 2's children who are most sensitive to context. Even here the coupling is weak, not absent, since susceptibility is a matter of degree (Belsky & Pluess, 2009). The child who sees the input and declines it is the reflective exit described below, which the section on the two perspectives calls I3 taken up: the pressure registers, is seen as pressure, and is chosen against. This is the only true exit of the three cases; the second is a weaker coupling, not an exit. The two have opposite signatures that a cohort can tell apart: weak transmission of everything in the low-susceptibility child, and strong transmission of the chosen content with weak transmission of the rest in the reflective one. One more case can look uncaptured and is not. The exempted adult can resemble the rebel or the reflective case and is neither: the rebel's conduct is the norm turned inside out, so it changes when the norm changes, while the exempted adult's follows its own gain, so it changes when

incentives change; and the reflective adult declines the norm for a content it has chosen and passes that on, while the exempted adult declines it for itself and keeps it for others. The exempted self is captured, by its own loop (Figure 1D), and its signatures are conduct that differs between watched and unwatched settings, for the hidden face, and conduct that follows incentives whether watched or not, for the open one. No one escapes the setting entirely on this reading, since the language, the sense of what is normal and the terms of any refusal all come from it; the model claims only that those who are not captured fall into one of these two cases.

2.1 Mixed settings: when family and wider environment differ (Measured and Interpretive)

The four loops in Figure 1 are the pure cases, in which family and wider norms point the same way. Families and wider settings can also differ, and the model treats them as two coupled loops that need not carry equal weight. Figure 2 draws three cases, two conflicts (A and B), in which both loops are present and pull opposite ways, and the one-sided case of the neglected child in a nurturing setting (C), and Table 2 sets out all the combinations with the pattern the model predicts for each. The predictions are hypotheses (Interpretive), not findings.



Figure 2. Three cases in which family and wider setting differ. Panel A: nurturing family in a wider setting whose norm permits harm. Panel B: coercive family in a nurturing wider setting. Panel C: absent or neglectful family in a nurturing wider setting, with the exit through another adult drawn. A and B are conflicts, with both inputs present; C is one-sided, with the family input weak (dotted). In each, the predicted pattern is shown at the foot. Coercive means harsh and inconsistent control; the consistent form and the remaining combinations are set out in Table 2. The predictions are hypotheses (Interpretive), tested by H2.

Absent or neglectful parenting is a further setting of the Parent input, not a kind of control: there is little to adapt to, so the family loop runs at low gain and the social superego, the norm of whoever is around, takes the Parent's place by default. The model predicts that the neglected child tracks the wider norm more closely than a parented child does, for better or worse, and that the exit through another adult matters most for this child (Table 2, last row). Measured. Neglect carries a raised risk of later antisocial conduct alongside abuse (Widom, 1989), and father absence, which is not neglect but a reduction of the family input, has effects on children's social-emotional adjustment that survive rigorous designs, though smaller than crude comparisons suggest (McLanahan et al., 2013). Interpretive. Absence changes the weights in the loop; it is not a new mechanism.

The neglected child in a nurturing setting (Figure 2C) is the most informative case the model has. The two conflicts depend on the relative weights of two loops, which are hard to predict and hard to test. In the one-sided case the family input is near zero, so whatever the child becomes is the setting's doing, and the exit through another adult stops being a side note and becomes the whole mechanism. Measured. In the Kauai Longitudinal Study, the one in three high-risk children who grew into competent adults all had at least one person who accepted them unconditionally, often a grandparent, an older sibling, a teacher or a neighbour (Werner, 1993). In Bucharest, young children in institutional care who were randomised to foster care showed cognitive recovery, largest in those placed youngest (Nelson et al., 2007). In a random-assignment evaluation of Big Brothers Big Sisters, mentored youths aged 10 to 16 were less likely over 18 months to start using drugs or alcohol, to hit someone or to skip school (Grossman & Tierney, 1998). Interpretive. Only Bucharest approaches the one-sided case, and what it shows is cognitive recovery, not the model's outcome; Kauai's high-risk children and the Big Brothers youths had parents, so those two are partial fits, in which an outside adult supplied what the family input lacked rather than replacing it. They are the nearest view of the horizontal channel the model offers, not a clean one, and the Bucharest timing effect is Elder's (1998) point about timing in loop form.

Figure 3 takes the neglected child through all three settings. The family input is the same empty box in each panel; only the content of the norm varies, and in each case a carrier from the setting, an adult, an institution or a gang, does the work the Parent would have done. Paper VI (section 4) reads a trained AI model as a further carrier. Measured. Order has a case as well as nurture, a partial one: among the Gluecks' delinquent boys from poor Boston homes, adult social bonds, a stable marriage and stable work, led to desistance beyond what the boys' childhood differences predicted (Sampson & Laub, 1993), and when the same men were traced in later life, marriage, meaningful work and military service, with its structure, discipline and sense of belonging, were the turning points that had taken them out of crime (Laub & Sampson, 2003); these men were not neglected children in the model's sense, and the turning points came in adulthood, so the case shows order supplied by the setting changing what a person passes on, not the cell's prediction of order adopted wholesale in childhood. Harm has its cases too, partial in the same way: In the barrios of southern California the gang served as parent, school and law in the absence of other role models (Vigil, 1988), and in the Rochester Youth Development Study the boys who joined gangs, about a third of the sample, offended violently more while they were members than before they joined or after they left, a facilitation effect that survived adjustment for other risk factors (Thornberry et al., 2003); neither study establishes that the boys were neglected, and Vigil's (1988) report of the gang serving as parent where other role models were absent is the nearest to the cell. Interpretive. These are partial fits to the three cells of Table 2's last row, with the mismatches stated; together they illustrate, rather than establish, the model's central claim: when the family is silent, the child becomes what the setting is. H2 therefore carries one more prediction: under neglect the outcome should track the norm's content more closely than under any parented condition, in all three directions, not only the buffering one.



Figure 3. The neglected child taken up by three settings. The same child with the same weak family input (dotted), taken up by a nurturing setting (Panel A), a setting whose norm is order (Panel B) and a setting whose norm permits harm (Panel C). In each, a carrier from the setting, another adult, an institution or a gang, does the work the Parent would have done, and the predicted pattern is shown at the foot with the measured cases named in the text. Only the content of the norm varies across the panels. The predictions are hypotheses (Interpretive), tested by H2.

Two features follow. The weaker loop can be overridden, so changing one input can change what is passed on even when the other stays the same. And the mixed cells are where the model can be shown wrong: if a nurturing wider setting makes no difference to children who are parented harshly, or nurturing parenting makes no difference in a controlling setting, the claim that the loops are coupled would fail. H2 (Table 3) tests this, and adds a prediction a model without the absent case cannot make: the norm's effect, in whichever direction, should be largest under absent or neglectful parenting, where the family loop offers least resistance; not a buffer, since there is little to buffer, but a larger main effect of the norm's content.

Table 2. Combinations of parenting and wider norms, and the pattern the model predicts for each.

Parenting	Nurturing norm	Norm of order	Harm-permitting norm
Nurturing (warmth, with limits)	Cyclic nurture (Figure 1C). Predicted: the most stable protective loop.	Buffered but contested. A nurturing Parent inside a rule-bound setting. Predicted: rule-following with self-direction kept; the balance depends on how strongly each loop reinforces, on the child's age, and on whether the family's values are shared outside the home.	Buffered but exposed (Figure 2A). The family forbids what the norm licenses. Predicted: transmission of harm reduced, most where another adult or group shares the family's values.
Controlling, consistent (authoritarian)	Compensated. Predicted: rule-following at home, self-direction supplied outside, and the order loop weakening with age.	Cyclic order (Figure 1B). Predicted: the most stable order loop.	The switch. Order with a licensed target (Altemeyer, 1981). Predicted: the main route from order into harm.
Controlling, coercive (harsh, inconsistent)	Compensated but incomplete (Figure 2B). A competing loop from teachers, peers, mentors or institutions. Predicted: reduced transmission, most in children more sensitive to context (Belsky & Pluess, 2009; Rutter, 1987; Werner & Smith, 1982).	Contained. Coercion learned at home but punished outside. Predicted: partial transmission, displaced into approved channels.	Cyclic harm (Figure 1A). Predicted: the most stable harmful loop.
Defecting or overvaluing (the rule kept for others, or warmth without limits; Figure 1D)	Defection contained. Predicted: limits supplied outside the home, applied to the child as to others, weaken the self-exemption with age; the hidden face where they are not.	Defection masked. Predicted: the hidden face, rules kept in public and broken in private, since open breach is punished; the child learns the mask with the rule.	Defection open: the second route into harm. Predicted: no mask is needed; where the exempted gain influence, defection merges into cyclic harm (H7).
Absent or neglectful	Filled from outside, best case (Figures 2C and 3A). Predicted: the wider norm takes the Parent's place; the exit through another adult matters most.	Filled from outside (Figure 3B). Predicted: order adopted wholesale, with little self-direction.	Filled from outside, worst case (Figure 3C). Predicted: a gang or street code as substitute family; the cell in which a change of norm should matter most (H2's neglect prediction).

Note. The predictions are hypotheses (Interpretive), not findings. Nurturing parenting includes firm limits given with warmth (Baumrind's authoritative pattern). A defecting parent keeps the rule for others and breaks it, hidden or openly; indulgent parenting is warmth without limits, with the child held above others (Brummelman et al., 2015; Maccoby & Martin, 1983). H1 tests the split between order and harm; H2 tests the buffering cells; H7 tests the defection row.

2.2 Related work and what is new (Measured and Interpretive)

None of the separate mechanisms here is new. Cycles of harm across generations are well documented: abused and neglected children are at raised risk of later violence (Widom, 1989), though most do not repeat it (Kaufman & Zigler, 1987). One of the best-documented transmissions is of attachment: in a meta-analysis of 18 samples, a parent's attachment state of mind predicted the infant's attachment, with 75 per cent agreement on secure versus insecure, while parental sensitivity carried only part of the link, a shortfall known as the transmission gap (van IJzendoorn, 1995). Social learning theory explains how behaviour is acquired by watching and being rewarded (Bandura, 1977). Coercion theory shows how harsh parenting and a child's resistance reinforce each other (Patterson, 1982). Parenting-style research separates controlling from supportive parenting (Baumrind, 1991). Norms research distinguishes what people do from what a group expects of them (Cialdini et al., 1990). Kohn (1969) found that parents whose work allowed little self-direction valued conformity to external authority in their children, parents whose work demanded self-direction valued it in theirs; the values were those each class wanted in its children, and the transmission itself

is inferred there, not measured. Tight cultures, with strong norms and punishment of deviance, arise under ecological and historical threat and are reproduced through narrow socialisation (Gelfand et al., 2011). Authoritarianism research describes the pro-order adult and traces it to punitive socialisation (Altemeyer, 1981; Duckitt, 2001). Experiments have shown moral hypocrisy, the appearance of fairness kept while serving oneself (Batson et al., 1997, 1999), and narcissism in children has been traced to parental overvaluation rather than to a lack of warmth (Brummelman et al., 2015). The idea that a society has its own superego goes back to Freud (1930/1961a), who wrote of a cultural superego.

What this article adds is integration and a framing, and it should be judged on that. First, it puts parenting and group norms into one loop with the same structure in a harmful, an ordering, a protective and a defecting form, so that either input can be changed, and it names two routes into harm, the switch from order and the exempted self. Second, it names the group-level input the social superego and defines it as the internalised in-group average, weighted by influence, which links the loop to the in-group paradox described below (section 3; paper VI, Table 2 in section 7, lists it among the failures that the two halves of peace interrupt) and which H7 puts to its first test. Third, it places gene-environment sensitivity as a modifier of the loop. Whether the integration adds anything beyond the existing theories is for the tests proposed below to show.

2.3 Positioning in family theory (Interpretive)

The model sits where several established family theories meet. Family systems theory describes a multigenerational transmission process by which patterns of relating pass down family lines, and it names the main protection as differentiation of self, the degree to which a person can keep the intellectual system distinct from the emotional system and act from thought rather than reactivity (Bowen, 1978). The vertical channel of the loop is a version of that process with its content specified. Differentiation, acting from thought rather than reactivity, is, in the terms of the section on the two perspectives, understanding setting the course for feeling within the family; the reflective exit described below adds the seeing that makes that direction a choice. Ecological systems theory places the family inside wider settings, from school and peer group to culture (Bronfenbrenner, 1979); in its terms the mixed cases of Table 2 are cases in which the family and the settings around it pull the child in different directions. Life course theory adds that lives are linked across generations and that the timing of an influence matters (Elder, 1998), which is why the child's age enters the predictions for the mixed cases. Kohn's (1969) work on class and conformity is the order loop seen from the sociology of the family: values shaped by the parent's work and passed to the child.

Contextual family therapy describes invisible loyalties that bind the generations and a revolving slate, in which revenge taken on a substitute creates a new victim, so that an injury done in one generation is claimed back from the next (Boszormenyi-Nagy & Spark, 1973). That is the loop's revenge form, the harm loop when its content is retaliation, inside the family. Symbolic interactionism supplies the closest relative of the social superego: the generalized other, the organised attitudes of the whole community that a person takes into their own conduct (Mead, 1934). The social superego adds one claim to Mead's concept, that the internalised expectation carries a felt penalty, guilt or shame, when it is broken. That penalty is not new in itself: Scheff (1988) made shame the emotion that enforces conformity to social expectation, and Bicchieri (2006) built an account of social norms on expectations about what others do and approve, backed by sanctions. The social superego differs from both in where it puts the norm, inside a loop that the family and the wider setting share and pass on, and in deriving it from the average of observed behaviour rather than from expectation alone; the fifth objection below sets out how to test whether the felt penalty is needed. Taken together, the model's contribution to these theories is to treat the family and the wider setting as two coupled loops with the same structure, to name the group-level input, and to state in advance what would count against the coupling.

2.4 Exits from the loop (Measured and Interpretive)

The loop names its own exits. The first is outside the person: a competing loop supplied by another adult, a school or a peer group (Table 2; Figure 2C). The second is inside: the capacity to see the loop while in it.

Measured. Attachment research has measured the second. Mothers who had experienced deprivation but were high in reflective functioning, the capacity to understand behaviour in terms of one's own and others' mental states, were much more likely to have securely attached infants (Fonagy et al., 1994). Some adults who were insecurely attached as infants are secure as adults and do well in close relationships (Roisman et al., 2002), so the pattern is not fixed.

Interpretive. In the model's terms, reflection turns the Adapted Child's automatic response into something that can be seen, and what can be seen can be chosen instead of passed on. Summerhayes (2026b) places this capacity with the witness, the self looking at its own perspectives, a term developed in Summerhayes (2026a) and set out in the section on the two perspectives below; differentiation of self is a near relative of it in family systems theory (Bowen, 1978). Contextual therapy describes a related release, exoneration, in which an adult's reassessment of a parent's own history lifts the burden of blame from the parent (Boszormenyi-Nagy & Krasner, 1986). For the order loop the exit is the same reflection applied to a rule: a rule seen as a rule can be kept or changed rather than obeyed. For the defection loop it is the same reflection turned on oneself, the self-exemption seen as an exemption. **Measured.** Self-awareness brought conduct into line with a moral standard only when the standard had been made salient beforehand; otherwise people adjusted the standard to fit the conduct (Batson et al., 1999). **Interpretive.** The third exit is forgiveness: in the empathy paper's sense, concern directed by understanding toward the offender (Summerhayes, 2026b; paper II, section 4), of which the measurable minimum used here is ceasing to retaliate. It ends the loop's revenge form (see the culture of honour, below). H6 (Table 3) tests whether reflective capacity weakens both channels, the vertical one through parenting and the horizontal one through norms.

3. The social superego and the in-group paradox (Interpretive)

Groups differ in the norms they enforce. Where an in-group's norms favour its members over outsiders, people who conform to the group's superego may be kind to members and hostile to out-groups. **Measured.** Group membership shapes behaviour toward in-group and out-group members, as social identity theory describes (Tajfel & Turner, 1979). **Interpretive.** The paradox is that the same mechanism that produces cohesion also produces division. It is a group-level relative of the humanity paradox of Summerhayes (2026b, Section 4), in which the capacity that lets people cooperate is the one that lets a few exploit the rest; here it is the norm that binds a group that sets it against others. Unity across groups would need a norm broad enough to cover the whole species, but each narrower identity pulls toward its own.

This is a hypothesis about how norms scale, not a measured result, and it is not among the seven tested here. A test would compare how strongly norms of in-group loyalty and of universal concern predict conduct toward out-groups; if loyalty norms predicted conduct toward out-groups no worse than universal-concern norms do, the mechanism would not be producing division and the paradox would not bite.

3.1 A measured case: the culture of honour (Measured and Interpretive)

Measured. In a set of experiments, young men from the US South who were insulted by a stranger showed, compared with men from the North, a rise in cortisol and testosterone, and were more likely to respond aggressively and dominantly (Cohen et al., 1996). Nisbett and Cohen (1996) explain the norm behind this, that an insult must be answered, as a legacy of herding settlers in regions where the law was too weak to protect livestock. Across the United States, historical settlement by the Scots-Irish predicts higher homicide today only in the South and, more generally, where institutions were

historically weak, and the effect has faded as institutions converged (Grosjean, 2014). These findings describe a norm, not the people who hold it; the same research shows it weakening as conditions changed.

Interpretive. This is the social superego in its plainest form: a group expectation, held as one's own, felt in the body, and passed on for generations after the conditions that made it pay had gone. It is also the loop in its revenge form. Each retaliation answers the last, so the cycle rebuilds itself, and it can pass from parent to child as a feud. Game theory shows the exit. Where mistakes happen, strict reciprocity (tit-for-tat) locks two players into endless retaliation, while a strategy that forgives a share of defections, generous tit-for-tat, breaks the cycle and is favoured in simulated populations (Nowak & Sigmund, 1992; paper II, section 5). Paper III (section 3, Table 1) gives the same release for the vendetta. Forgiveness is therefore the exit from the loop's revenge form, as reflection is the exit from its compliance forms, the order and harm loops when their content is obedience or coercion rather than retaliation. In the loops' terms the culture of honour is the switch made into a rule: an order whose rule licenses a target, whoever gives the insult, so that retaliation is order's form and harm's content at once, which is why it sits on the switch between the order loop and the harm loop. And the fading of the effect where institutions strengthened is a measured case of what Table 2 predicts: a change in the wider setting changes what is passed on.

3.2 A measured case: tightness and authoritarianism (Measured and Interpretive)

Measured. Across 33 nations, cultures with strong norms and low tolerance of deviance, tight cultures, were more common where ecological and historical threat was higher, from population density and resource scarcity to disease and territorial conflict, and tight nations had more autocratic governments, more police per capita, stricter punishments and lower crime rates (Gelfand et al., 2011). Within one country, parents whose work allowed little self-direction valued conformity to external authority in their children, and parents whose work demanded self-direction valued it in theirs, the values of each class being those its parents wanted in their children (Kohn, 1969). In Duckitt's (2001) dual-process model, right-wing authoritarianism, measured by Altemeyer's (1981) scale, is linked in cross-sectional paths to punitive socialisation and a view of the world as dangerous.

Interpretive. This is cyclic order measured in three settings, the nation, the class and the family: a norm of order, reproduced through narrow socialisation, that pays under the threats that produced it and persists after them. Its exit is reflection applied to a rule rather than a parent, and its risk is the switch described in the model section, aggression in the name of authority, which is the in-group paradox above seen from inside one person.

3.3 A measured case: defection, hidden and open (Measured and Interpretive)

Measured. In three small studies of female students, 16 of 20 who could assign a pleasant task to themselves or to another participant took it themselves, and only one later called that the moral choice; when a coin was offered as a fair way to decide, half flipped it, and 9 of the 10 who flipped still gave themselves the pleasant task (Batson et al., 1997). In five studies, condemning a wrong made a person look virtuous, sometimes more than a direct claim of good conduct did, and people who condemned what they did themselves were judged more harshly than people who lied outright about it (Jordan et al., 2017). Pedestrians crossed against a signal more often when a model did, and more often still when the model wore high-status clothes (Lefkowitz et al., 1955). In six field experiments, signs that others had broken a norm or rule made people more likely to break other norms and rules (Keizer et al., 2008); a replication of the designs found the same direction for littering and small theft, with the effect limited to breaches that cost little (Keuschnigg & Wolbring, 2015). Across 23 countries, young people from societies with more corruption, tax evasion and political fraud were less honest in an anonymous task run years after the societal measures were taken (Gächter & Schulz, 2016); the link is correlational.

Interpretive. These are the parts of cyclic defection, not a test of it; the two routes in are given with the loop above, the modelling route in one sitting (Hays & Carver, 2014) and across years,

retrospectively (Setoh et al., 2020). The participants of Batson et al. (1997) show the hidden face, and the later finding given with the exits above shows that the mask can deceive its wearer as well as others. Jordan et al. (2017) show why the mask works: condemnation signals the virtue it hides. Lefkowitz et al. (1955) show a breach weighted by status, and Keizer et al. (2008) and Gächter and Schulz (2016) show breaches spreading, and honesty lower where they are common. None shows the loop: that self-exemption is learned from a parent and passed on, or that the breaches of the influential move the norm more than the same breaches by others. H7 tests both.

3.4 Felt and understood: the social superego in two perspectives (Interpretive)

Summerhayes (2026a, 2026b; papers I and II) reads experience through three perspectives: the felt perspective (I1), the understanding perspective (I2) and the witness's eye (I3), through which the other two are seen together; the eye does not act, and the witness is the self looking through it. In those terms the social superego is a structure held in the understanding perspective: the norm is known, as what the group does and so expects. Its force comes through the felt perspective, as guilt or shame at breaking it. The superego therefore has both perspectives, as the ego does, so neither is fundamental: both are structures built from the felt and the understanding perspectives. This is the distinction that H1 measures, the perceived norm and the person's own guilt, and that the reply to the fifth objection below relies on. Mead's (1934) generalized other is, on this reading, the understanding half of the social superego. Internalisation is taken to run through the same observation and reward (Bandura, 1977), with the penalty then applied by the person to themselves. The order loop's superego is the rule known and guilt at breaking it; the harm loop's superego licenses the harm, so no guilt follows it; and the switch between the loops is whether the felt penalty attaches to doing the harm or to failing to do it. The defection loop's superego is known and not felt, at least for oneself: the penalty is applied to others' breaches and not to one's own. That is understanding serving feeling at its negative pole, the empathy paper's exploitation (Summerhayes, 2026b; paper II, section 4), held behind a norm rather than licensed by one.

Read this way, the move from Free Child to Adapted Child is one route by which the ego forms, under pressure; in the nurture loop the ego forms too, through limits given with warmth, so that the Free Child develops rather than complies; in the defection loop it forms around a rule held for others, learned by example or by exemption; and the difference is in what is internalised, not in whether anything is. With the ego comes the end of a child's first innocence, which Summerhayes (2026b) places with the first self-conscious emotions; the loop is how that formation is passed on. The reflective exit described above is, in the same reading, I3 taken up: seeing the loop from outside it while still in it, of which reflective functioning and psychological flexibility are the nearest measured correlates (Fonagy et al., 1994; Hayes et al., 2006). On this reading innocence is not lost for good: a second innocence, with the edge of the ego seen rather than removed, is found when I3 is taken up. These readings are offered as a framing, and none of the hypotheses depends on them.

4. A worked example: MAOA and childhood adversity (Measured and Interpretive)

Paper V of the series is the protocol for a systematic review and meta-analysis of this interaction, written to test it.

4.1 What is known (Measured)

The monoamine oxidases A and B (MAO-A, MAO-B) differ in what they break down: MAO-A mainly serotonin and noradrenaline, MAO-B other amines such as phenylethylamine, and both handle dopamine. The MAOA gene is on the X chromosome. A rare loss-of-function mutation in MAOA, described in one Dutch family, was associated with impulsive aggression in affected males (Brunner et al., 1993). A common variant, the uVNTR promoter polymorphism, affects transcriptional activity in vitro (Sabol et al., 1998). In a birth cohort, boys with the low-activity variant who were maltreated showed more antisocial outcomes than maltreated boys with the high-activity variant (Caspi et al.,

2002). Two meta-analyses found the interaction in males, with small effects (Byrd & Manuck, 2014; Kim-Cohen et al., 2006).

4.2 What is not known (Measured and Interpretive)

The interaction does not replicate in every study, and effects in females are less clear. The result is a probabilistic risk modifier, not a cause of violence. The two pooled tests cited here (Byrd & Manuck, 2014; Kim-Cohen et al., 2006) are the most recent I found in a limited, free-source search on 13 October 2026, and a new pooled test, for which a protocol is in preparation (Summerhayes, 2026e), would be needed to settle the evidence behind H3. Candidate-gene-by-environment studies as a class have been criticised for small samples and poor replication (Duncan & Keller, 2011). H3 is therefore the weakest hypothesis here and is offered as a test, not as a pillar of the model.

Because the gene sits on the X chromosome, males carry one copy. Where the low-activity allele has frequency q , affected males outnumber homozygous affected females by about $1/q$; there is no fixed ratio. DNA methylation of MAOA correlates with MAO-A activity in the brain of healthy men (Shumay et al., 2012). That is an association within individuals; it does not show that social environments rewrite the gene.

4.3 How it connects to the model (Interpretive)

The loop model predicts that sensitivity to controlling or nurturing input is not the same for everyone. MAOA is a candidate modifier. Its use here is illustrative: gene-environment interaction is one route by which the loop might bite harder in some people than in others. It is not evidence for the loop; its link to the loop is that, if the interaction holds, the harm loop's vertical channel bites harder in carriers of the low-activity allele, so H3 tests a modifier of the loop's gain and not the loop. The same alleles may also matter on the supportive side. Differential susceptibility predicts that allele carriers do better, not only worse, under supportive conditions (Belsky & Pluess, 2009); the evidence for this is limited, so it is treated here as a hypothesis, which allele carriers doing no better, or worse, than non-carriers under supportive conditions would count against, and it is the idea behind the compensated cell in Table 2. The same difference in sensitivity has a lower end: the child least sensitive to context is the least shaped by either loop, for harm or for nurture, as the model section notes.

Two cautions on inheritance. First, rodent work shows that maternal behaviour can change gene regulation (Weaver et al., 2004), but transmission of epigenetic marks across generations in humans is unproven and contested (Heard & Martienssen, 2014). The vertical channel in the model is therefore parenting and learning, with epigenetic inheritance only a hypothesis to be tested. Second, nothing here supports ranking people by genotype; variants are common and their effects small.

5. Dark Tetrad traits and empathy (Measured and Interpretive)

Measured. The Dark Triad of narcissism, Machiavellianism and psychopathy (Paulhus & Williams, 2002) became a Tetrad when everyday sadism was added (Paulhus, 2014). Empathy has at least two parts. Psychopathy has been associated with impaired emotional (affective) empathy, with cognitive empathy often reported as relatively preserved (Blair, 2005). A later meta-analysis found the link between psychopathy and reduced emotional empathy to be clearer than the link with cognitive empathy, which was weaker (r about $-.23$, against about $-.35$ for affective empathy) and varied widely across studies (Shukla & Upadhyay, 2025). Cognitive and emotional empathy rely on partly separate brain regions (Shamay-Tsoory et al., 2009), and compassion, as distinct from empathy, has its own network (Singer & Klimecki, 2014).

Interpretive. If compassion needs both forms, a person with adequate cognitive empathy and low emotional empathy might read others accurately without being moved by them. The empathy paper (Summerhayes, 2026b; paper II, sections 4 and 6) separates two profiles that share that pattern: without a self-serving drive it is indifference; with one, understanding in the service of feeling at its negative pole, it is the harm profile, the few whom that paper's humanity paradox describes. In a

system that rewards status and profit, the second may do well, and a loop that rewards the traits could make them more common in positions of influence; the defection loop (Figure 1D) is how such a loop would run, through the weighting by influence that H7 tests. Duckitt's (2001) second path, linking unaffectionate socialisation to social dominance orientation in cross-sectional data, points toward the harm profile; the order loop, if it shapes the pattern at all, would be expected to produce indifference. This is a hypothesis, and H4 (Table 3) tests the loop's part in both. It does not say that anyone who holds power, or any named group or institution, has these traits.

Summerhayes (2026b; paper II, section 7) develops this reading into a test of its own. The question for this article is different: whether the loop helps produce the profiles. Measured. In an adoption study of 561 families, severe antisocial behaviour in birth mothers, who had little or no contact with their children, predicted callous-unemotional behaviour at 27 months, and adoptive mothers' positive reinforcement buffered that heritable risk (Hyde et al., 2016). In a cross-sectional study of 372 non-clinical adults, recalled psychological aggression from parents was uniquely associated with psychopathy and sadism, and recalled severe assault with Machiavellianism, narcissism and psychopathy (Galan et al., 2025). Interpretive. Hyde et al. show that parenting moderates a heritable pathway. That is the kind of interaction the loop needs, and also why shared genes must be controlled (the fourth objection below). H4 asks whether the horizontal channel, controlling norms, adds to the pattern.

6. Testable hypotheses (Interpretive, with proposed tests)

Table 3 lists seven hypotheses, each with a design and a result that would count against it. H1 tests the two inputs and the split between order and harm, H2 the coupling between family and setting, H3 the genetic modifier, H4 the loop's part in the empathy profiles, H5 the added value of changing norms, H6 the reflective exit, and H7 the defection loop and the weighting by influence. The interaction hypotheses need large samples. For H1 and H2, a small interaction (f^2 of 0.02) needs about 400 participants at 80 per cent power and a 5 per cent two-sided error rate if measured without error, and several times that with retrospective self-report, which weakens interaction terms, so both are proposed for cohorts of several thousand. For H4 the harm profile is rare in community samples, so the cohort must either be screened and enriched for callous traits, with estimates weighted back, or be large enough to hold a few hundred people in each profile. For H6 each channel's term is a small interaction of the same order and needs a sample of the size H1 does. H5 is powered for its interaction in the trial design itself. H7's first part is an increment over warmth in a cohort, which needs less than an interaction does, and its second an experimental contrast powered in its own design.

Table 3. Testable hypotheses, proposed designs and results that would count against each.

No.	Hypothesis	Design	Would count against it
H1	Retrospective reports of coercive parenting and a harm-permitting local norm each predict adult antisocial conduct independently and, if the loops are coupled, more strongly together than their sum (an interaction); reports of consistent control and a norm of order instead predict rule-following and authoritarian attitudes; and the person's own guilt or shame at breaking the norm adds to the two perceived norms in predicting conduct.	Cohort or cross-sectional survey, with parenting measures that score care and control separately (for example the Parental Bonding Instrument; Parker et al., 1979) and separate consistent rule-setting from harsh and inconsistent discipline (for example the Alabama Parenting Questionnaire; Shelton et al., 1996, which is validated for parents and children and would need adapting and validating for adults' retrospective reports); perceived descriptive and injunctive norms and their content (what people around the respondent do and expect, and whether that licenses harm to anyone); the person's own guilt or shame at breaking them, measured with scenario items about hypothetical breaches rather than the respondent's own conduct, and in a cohort before the conduct measure; a validated antisocial behaviour scale or official records; and an authoritarianism scale (Altemeyer, 1981). A genetically informative design (adoption or children-of-twins) is needed to separate environmental transmission from shared genes; a sibling comparison cannot, since parenting style and local norm are shared by siblings.	In the survey design: no independent or joint effect after adjusting for socioeconomic status; consistent and coercive control predict the same outcomes, with norm content adding nothing; or guilt adds nothing to the two perceived norms. In the genetically informative design: the association disappears once shared genes are controlled
H2	Nurturing parenting buffers the effect of controlling norms and nurturing norms buffer the effect of controlling parenting (interactions), and under absent or neglectful parenting the outcome tracks the norm's content more closely than under any parented condition (a main effect of the norm's content that is larger in the neglected group; Table 2; Figure 3).	Same data; test interaction terms, with absent or neglectful parenting as a separate category and the norm's content (nurturing, order, harm-permitting) coded	Additive effects only; or the norm's main effect is no larger under neglect than under parented conditions
H3	In males, the effect of childhood maltreatment or harsh parenting on antisocial outcomes is larger for the low-activity MAOA variant than for the high-activity variant; whether the interaction extends to milder controlling parenting is a secondary question.	Pre-registered analysis in a large cohort or consortium, with the allele coding and the threshold for a meaningful interaction fixed in advance	Interaction absent in the pre-registered sample, or present only in the unadjusted analysis

No.	Hypothesis	Design	Would count against it
H4	Coercive parenting and harm-permitting local norms each predict, in adulthood, the harm profile of the empathy paper (Summerhayes, 2026b): lower emotional empathy with cognitive empathy preserved and a self-serving or dominance drive, beyond heritable risk; consistent control and norms of order predict instead the same empathy pattern without the drive, which that paper calls indifference; and nurturing parenting buffers both.	Genetically informative cohort (adoption or children-of-twins) with parenting and norm measures, an empathy scale that scores emotional and cognitive empathy separately, the Inventory of Callous-Unemotional Traits (Frick, 2004) and a measure of proactive aggression (Raine et al., 2006), which defines the drive; each profile defined in advance as a conjunction of thresholds on standardised scores (for example, emotional empathy at least 1 <i>SD</i> below the mean, cognitive empathy within 0.5 <i>SD</i> of it, and proactive aggression at least 1 <i>SD</i> above it for the harm profile and below the mean for indifference)	No association once heritable risk is controlled; the loop lowers both kinds of empathy equally, with no specific profile; or coercive and consistent control predict the same profile, with or without the drive
H5	Parenting programmes that shift parents from controlling to nurturing styles are known to reduce child conduct problems in trials (Furlong et al., 2012); the new prediction is that the effect is larger and lasts longer where local norms also change.	Randomised or cluster-randomised trials that vary the norm component (for example school or community norms) as well as the parenting component, powered for the interaction, with a manipulation check on the change in norms and follow-up over years	No larger or longer effect where norms also change
H6	Reflective capacity (reflective functioning or psychological flexibility) weakens the association between controlling parenting received and controlling parenting given, and between the perceived controlling norm and conformity to it (moderation on each channel).	Cohort of parents with retrospective reports of parenting received, a validated measure of reflective capacity, observed parenting of their own children, perceived-norm measures and a measure of conformity to the norm (conduct or endorsement), with follow-up of their children. Each channel is tested as an interaction term, parenting received by reflective capacity and perceived norm by reflective capacity. The existing support (Fonagy et al., 1994) is for infant attachment security, not for less controlling parenting, so it does not stand in for the parenting-channel test	A null interaction term on a channel, which fails the hypothesis on that channel; or reflective capacity strengthening the association rather than weakening it
H7	Parental defection (the rule kept for others and broken by the parent) and parental overvaluation (warmth without limits, with the child held above others) each predict adult self-exemption beyond parental warmth, and most strongly together: conduct that differs between watched and unwatched settings (the hidden face) or follows incentives whether watched or not (the open face); and a breach by a person of high status that goes unsanctioned moves the conduct and descriptive norm of those who see it more than the same breach by a person of low status (the weighting by influence).	(a) Cohort with reports of the parent's own rule-breaking and double standards, and parent reports of overvaluation and warmth in childhood (as measured by Brummelman et al., 2015), and, in adulthood, conduct measured both when the person believes it is observed and when it is anonymous, with the incentive varied (for example a private coin-flip assignment, Batson et al., 1997, or an anonymous die roll, Gächter & Schulz, 2016); a genetically informative design to separate parenting from shared genes. (b) Field or laboratory experiments in which a confederate of high or low status breaks a local rule, sanctioned or not (as in Keizer et al., 2008; Lefkowitz et al., 1955), measuring observers' later breaches and their reported descriptive norm	(a) Neither parental defection nor overvaluation adds to warmth in predicting either face; (b) status makes no difference to how far an unsanctioned breach moves observers' conduct or norm, or sanctioned and unsanctioned breaches move it equally

7. Objections and replies (Interpretive)

The first objection is that the model is too general to be wrong. Every hypothesis in Table 3 names a result that would count against it, and H3, H4, H6 and H7 can fail in a single study.

A second objection is that ego states are not scientific. They are used here as descriptive labels, and the hypotheses use validated measures of parenting, norms and conduct instead.

A third is that the MAOA evidence is mixed. I agree, and the section on MAOA says so. MAOA is an illustration of gene-environment interaction, not an essential part of the model.

A fourth is that parenting and behaviour could be linked through shared genes, not through the loop. This is a real confound, because parents also pass on genes, and children also shape how their parents behave. H1 and H4 therefore call for genetically informative designs, and H5 for randomised trials, which break that link.

A fifth is that the social superego is only a descriptive norm under another name. The difference claimed is internalisation. A norm that is only perceived is a fact about the group, while a superego norm is one the person punishes themselves for breaking (guilt or shame). H1 can test this by measuring the perceived descriptive norm (what people around you do), the perceived injunctive norm (what they expect) and the person's own guilt response separately; if guilt adds nothing to the two perceived norms, the construct is not needed.

A sixth is that the empathy profile is already the subject of other work (Summerhayes, 2026b). That work asks what the states of empathy and compassion are; this article asks how harm and care are passed on. Where the two meet, H4 here is about transmission, not about the profile itself.

A seventh is that the order loop is authoritarianism research under another name. Duckitt's (2001) model already runs from punitive socialisation to authoritarianism. What the loop adds is the coupling of that path with the harm path in one structure, the family-and-society coupling that lets either input be changed, and the single switch, a licensed target, that H1 and the authoritarian-aggression component can test.

An eighth is that defection, as used here, is narcissism, moral hypocrisy or game theory's defection under another name. Each is measured or modelled, as a trait, as conduct in a single setting and as a move in a game, and each is used here. What the loop adds is transmission: two inputs in the family, a parent who defects and a child held above others; two channels by which self-exemption is passed on; and the claim that, through the weighting by influence, it becomes a route into harm. H7 tests the input and the weighting, and if neither parental defection nor overvaluation adds to warmth the defection loop has no input of its own.

The last is that the account could stigmatise groups. The model names no individual and makes no claim of its own about any group, and it does not say that any person of status defects. The regional examples, the culture of honour and the barrios, report published findings about norms, and the honour findings show the norm fading as conditions changed.

8. Discussion (Interpretive)

8.1 Limits and risks (Experienced and Interpretive)

The model comes from my own observation and self-examination, and no new data are reported. Parent and child share genes, so the association between parenting and outcomes cannot be read as proof of the loop without genetically informative designs. The MAOA example rests on candidate-gene studies with mixed replication. Ego-state labels are descriptive, and the social superego is a construct whose added value over existing norm theory is still to be shown.

Labelling people as Dark Tetrad, as defectors, or as having particular genotypes, risks stigma; the model describes loops and conditions, not types of people. Calling norms controlling or nurturing is a

judgement that depends on context and culture, and the article gives no way to rank cultures. Forgiveness, as measured here, means at least ceasing to retaliate; it does not mean reconciling with, or staying close to, someone who remains a danger. Nothing here is health advice, and the one practice mentioned, mindfulness, is offered only as untested. I write as an independent researcher with no clinical role, and the model should be treated as a hypothesis until tested.

8.2 Practices that follow (Interpretive)

If the loop is real, the most direct levers are the inputs: parenting and norms. Evidence-based parenting support, and schools and workplaces that reward cooperation over fear, are consistent with the model. Reflective practices such as mindfulness may help individuals notice when they are enacting the loop; their effects on this loop have not been tested and should not be assumed. Reflective functioning, by contrast, has a measured link to secure attachment in infants of mothers who had experienced deprivation (Fonagy et al., 1994), and building it in parents is a direct way to test H6. Where the loop has taken its revenge form, the lever is forgiveness in the sense given above: an end to retaliation, not a return to danger. Where it has taken its order form, the lever is the content of the rule: order kept without a licensed target. Where it has taken its defection form, the lever is sanction that falls on the influential as it falls on others, so that the hidden face gains nothing by its mask and the open face is not copied, and, in the family, parents who keep the rules they set and do not hold the child above them. And since the social superego is weighted by influence, whom a person gives weight to is itself a lever: choosing the adults, peers and models one attends to is choosing the average one internalises, the exit through another adult taken up by choice.

8.3 What would change the model (Interpretive)

If H1 fails once shared genes are controlled, the loop is mostly inheritance and not transmission by parenting and norms. If consistent and coercive control predict the same outcomes, or the content of the norm adds nothing, the order loop collapses into the harm loop and the distinction should be dropped. If the perceived norm predicts conduct as well without a guilt component, the social superego adds nothing and should be dropped. If H5 shows no extra effect where norms also change, in a trial powered for the interaction and with the change in norms confirmed by a manipulation check, the case for a group-level input is not supported. If H2 shows additive effects only, so that neither a nurturing family nor a nurturing setting buffers the other, the claim that the two loops are coupled fails and the model reduces to two independent influences. If the pre-registered test of H3 is null, the MAOA illustration should be removed; the loop model does not depend on it. If H4 finds no association once heritable risk is controlled, the loop does not shape the empathy profiles, and the argument on Dark Tetrad traits rests on Summerhayes (2026b) alone. If H6 finds moderation on neither channel, reflection is not the exit the model claims; if on the parenting channel only, it is an exit from the family loop alone. If H7's first part fails, and neither parental defection nor overvaluation adds to warmth in predicting either face of defection, the defection loop has no input of its own and should be folded into the harm loop as one of its contents; if its second part fails, and status makes no difference to how far an unsanctioned breach moves the norm, the weighting by influence in the definition of the social superego has no support and should be dropped.

9. Conclusion

The idea is simple: harm, order, nurture and defection all repeat, and the repeating has parts that can be named, measured and changed. The model gives them names and puts parenting and group norms into one loop that can run toward harm, order, nurture or defection, with two routes into harm: a licensed target from order, and an exempted self that gains influence without sanction. Its contribution is integration and a framing, not a new mechanism.

H1 and H2 can be tested first, in existing cohorts with parenting measures, once items on the perceived norm and its content, including whom it licenses as a target, are added, since few cohorts have them; H4 and H6 can follow in existing adoption and parent cohorts. H5 would need a trial that varies the norm component. H7's experimental part can be run now, and its cohort part wherever

parental defection or overvaluation was measured in childhood. H3 should be tried only in a large, pre-registered sample. Whether the model is right depends on tests that others can run.

References

- Altemeyer, B. (1981). *Right-wing authoritarianism*. University of Manitoba Press.
- Anthropic. (2026). *Claude* (Sonnet 5.5, Opus 5.5, Haiku 5.5 and Fable 5.1 versions) [Large language model]. <https://claude.ai>
- Bandura, A. (1977). *Social learning theory*. Prentice-Hall.
- Batson, C. D., Kobrynowicz, D., Dinnerstein, J. L., Kampf, H. C., & Wilson, A. D. (1997). In a very different voice: Unmasking moral hypocrisy. *Journal of Personality and Social Psychology*, 72(6), 1335–1348. <https://doi.org/10.1037/0022-3514.72.6.1335>
- Batson, C. D., Thompson, E. R., Seufferling, G., Whitney, H., & Strongman, J. A. (1999). Moral hypocrisy: Appearing moral to oneself without being so. *Journal of Personality and Social Psychology*, 77(3), 525–537. <https://doi.org/10.1037/0022-3514.77.3.525>
- Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence*, 11(1), 56–95. <https://doi.org/10.1177/0272431691111004>
- Belsky, J., & Pluess, M. (2009). Beyond diathesis stress: Differential susceptibility to environmental influences. *Psychological Bulletin*, 135(6), 885–908. <https://doi.org/10.1037/a0017376>
- Berne, E. (1961). *Transactional analysis in psychotherapy*. Grove Press.
- Berne, E. (1964). *Games people play*. Grove Press.
- Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.
- Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, 14(4), 698–718. <https://doi.org/10.1016/j.concog.2005.06.004>
- Boszormenyi-Nagy, I., & Krasner, B. R. (1986). *Between give and take: A clinical guide to contextual therapy*. Brunner/Mazel.
- Boszormenyi-Nagy, I., & Spark, G. M. (1973). *Invisible loyalties: Reciprocity in intergenerational family therapy*. Harper & Row.
- Bowen, M. (1978). *Family therapy in clinical practice*. Jason Aronson.
- Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press.
- Brummelman, E., Thomaes, S., Nelemans, S. A., Orobio de Castro, B., Overbeek, G., & Bushman, B. J. (2015). Origins of narcissism in children. *Proceedings of the National Academy of Sciences*, 112(12), 3659–3662. <https://doi.org/10.1073/pnas.1420870112>
- Brunner, H. G., Nelen, M., Breakefield, X. O., Ropers, H. H., & van Oost, B. A. (1993). Abnormal behavior associated with a point mutation in the structural gene for monoamine oxidase A. *Science*, 262(5133), 578–580. <https://doi.org/10.1126/science.8211186>
- Byrd, A. L., & Manuck, S. B. (2014). MAOA, childhood maltreatment, and antisocial behavior: Meta-analysis of a gene-environment interaction. *Biological Psychiatry*, 75(1), 9–17. <https://doi.org/10.1016/j.biopsych.2013.05.004>
- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A., & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297(5582), 851–854. <https://doi.org/10.1126/science.1072290>
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6), 1015–1026. <https://doi.org/10.1037/0022-3514.58.6.1015>
- Cohen, D., Nisbett, R. E., Bowdle, B. F., & Schwarz, N. (1996). Insult, aggression, and the southern culture of honor: An “experimental ethnography.” *Journal of Personality and Social Psychology*, 70(5), 945–960. <https://doi.org/10.1037/0022-3514.70.5.945>
- Dawkins, R. (1982). *The extended phenotype: The gene as the unit of selection*. W. H. Freeman.

- Duckitt, J. (2001). A dual-process cognitive-motivational theory of ideology and prejudice. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 33, pp. 41–113). Academic Press.
[https://doi.org/10.1016/S0065-2601\(01\)80004-6](https://doi.org/10.1016/S0065-2601(01)80004-6)
- Duncan, L. E., & Keller, M. C. (2011). A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *American Journal of Psychiatry*, 168(10), 1041–1049.
<https://doi.org/10.1176/appi.ajp.2011.11020191>
- Elder, G. H., Jr. (1998). The life course as developmental theory. *Child Development*, 69(1), 1–12.
<https://doi.org/10.2307/1132065>
- Fonagy, P., Steele, M., Steele, H., Higgitt, A., & Target, M. (1994). The Emanuel Miller Memorial Lecture 1992: The theory and practice of resilience. *Journal of Child Psychology and Psychiatry*, 35(2), 231–257.
<https://doi.org/10.1111/j.1469-7610.1994.tb01160.x>
- Freud, S. (1961a). Civilization and its discontents. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 21, pp. 64–145). Hogarth Press. (Original work published 1930)
- Freud, S. (1961b). The ego and the id. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 19, pp. 1–66). Hogarth Press. (Original work published 1923)
- Frick, P. J. (2004). *Inventory of Callous–Unemotional Traits* [Unpublished rating scale]. University of New Orleans.
- Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database of Systematic Reviews*, (2), Article CD008225.
<https://doi.org/10.1002/14651858.CD008225.pub2>
- Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature*, 531(7595), 496–499. <https://doi.org/10.1038/nature17160>
- Galan, M., Pineda, D., Rico-Bordera, P., Piqueras, J. A., & Muris, P. (2025). Dark childhood, dark personality: Relations between experiences of child abuse and dark tetrad traits. *Personality and Individual Differences*, 238, Article 113089. <https://doi.org/10.1016/j.paid.2025.113089>
- Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., Aycan, Z., Boehnke, K., Boski, P., Cabecinhas, R., Chan, D., Chhokar, J., D'Amato, A., Ferrer, M., Fischlmayr, I. C., ... Yamaguchi, S. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, 332(6033), 1100–1104. <https://doi.org/10.1126/science.1197754>
- Grosjean, P. (2014). A history of violence: The culture of honor and homicide in the US South. *Journal of the European Economic Association*, 12(5), 1285–1316. <https://doi.org/10.1111/jeea.12096>
- Grossman, J. B., & Tierney, J. P. (1998). Does mentoring work? An impact study of the Big Brothers Big Sisters program. *Evaluation Review*, 22(3), 403–426. <https://doi.org/10.1177/0193841X9802200304>
- Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and commitment therapy: Model, processes and outcomes. *Behaviour Research and Therapy*, 44(1), 1–25.
<https://doi.org/10.1016/j.brat.2005.06.006>
- Hays, C., & Carver, L. (2014). Follow the liar: The effects of adult lies on children's honesty. *Developmental Science*, 17(6), 977–983. <https://doi.org/10.1111/desc.12171>
- Heard, E., & Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell*, 157(1), 95–109. <https://doi.org/10.1016/j.cell.2014.02.045>
- Hyde, L. W., Waller, R., Trentacosta, C. J., Shaw, D. S., Neiderhiser, J. M., Ganiban, J. M., Reiss, D., & Leve, L. D. (2016). Heritable and nonheritable pathways to early callous-unemotional behaviors. *American Journal of Psychiatry*, 173(9), 903–910. <https://doi.org/10.1176/appi.ajp.2016.15111381>
- Jordan, J. J., Sommers, R., Bloom, P., & Rand, D. G. (2017). Why do we hate hypocrites? Evidence for a theory of false signaling. *Psychological Science*, 28(3), 356–368. <https://doi.org/10.1177/0956797616685771>
- Kaufman, J., & Zigler, E. (1987). Do abused children become abusive parents? *American Journal of Orthopsychiatry*, 57(2), 186–192. <https://doi.org/10.1111/j.1939-0025.1987.tb03528.x>
- Keizer, K., Lindenberg, S., & Steg, L. (2008). The spreading of disorder. *Science*, 322, 1681–1685.
<https://doi.org/10.1126/science.1161405>

- Keuschnigg, M., & Wolbring, T. (2015). Disorder, social capital, and norm violation: Three field experiments on the broken windows thesis. *Rationality and Society*, 27(1), 96–126. <https://doi.org/10.1177/1043463114561749>
- Kim-Cohen, J., Caspi, A., Taylor, A., Williams, B., Newcombe, R., Craig, I. W., & Moffitt, T. E. (2006). MAOA, maltreatment, and gene-environment interaction predicting children's mental health: New evidence and a meta-analysis. *Molecular Psychiatry*, 11(10), 903–913. <https://doi.org/10.1038/sj.mp.4001851>
- Kohn, M. L. (1969). *Class and conformity: A study in values*. Dorsey Press.
- Laub, J. H., & Sampson, R. J. (2003). *Shared beginnings, divergent lives: Delinquent boys to age 70*. Harvard University Press.
- Lefkowitz, M., Blake, R. R., & Mouton, J. S. (1955). Status factors in pedestrian violation of traffic signals. *Journal of Abnormal and Social Psychology*, 51, 704–706. <https://doi.org/10.1037/h0042000>
- Maccoby, E. E., & Martin, J. A. (1983). Socialization in the context of the family: Parent-child interaction. In P. H. Mussen (Series Ed.) & E. M. Hetherington (Vol. Ed.), *Handbook of child psychology: Vol. 4. Socialization, personality, and social development* (4th ed., pp. 1–102). Wiley.
- McLanahan, S., Tach, L., & Schneider, D. (2013). The causal effects of father absence. *Annual Review of Sociology*, 39, 399–427. <https://doi.org/10.1146/annurev-soc-071312-145704>
- Mead, G. H. (1934). *Mind, self, and society: From the standpoint of a social behaviorist*. University of Chicago Press.
- Nelson, C. A., III, Zeanah, C. H., Fox, N. A., Marshall, P. J., Smyke, A. T., & Guthrie, D. (2007). Cognitive recovery in socially deprived young children: The Bucharest Early Intervention Project. *Science*, 318(5858), 1937–1940. <https://doi.org/10.1126/science.1143921>
- Nisbett, R. E., & Cohen, D. (1996). *Culture of honor: The psychology of violence in the South*. Westview Press.
- Nowak, M. A., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, 355(6357), 250–253. <https://doi.org/10.1038/355250a0>
- Parker, G., Tupling, H., & Brown, L. B. (1979). A parental bonding instrument. *British Journal of Medical Psychology*, 52(1), 1–10. <https://doi.org/10.1111/j.2044-8341.1979.tb02487.x>
- Patterson, G. R. (1982). *Coercive family process*. Castalia.
- Paulhus, D. L. (2014). Toward a taxonomy of dark personalities. *Current Directions in Psychological Science*, 23(6), 421–426. <https://doi.org/10.1177/0963721414547737>
- Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, 36(6), 556–563. [https://doi.org/10.1016/S0092-6566\(02\)00505-6](https://doi.org/10.1016/S0092-6566(02)00505-6)
- Raine, A., Dodge, K., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., Stouthamer-Loeber, M., & Liu, J. (2006). The Reactive-Proactive Aggression Questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. *Aggressive Behavior*, 32(2), 159–171. <https://doi.org/10.1002/ab.20115>
- Roisman, G. I., Padrón, E., Sroufe, L. A., & Egeland, B. (2002). Earned-secure attachment status in retrospect and prospect. *Child Development*, 73(4), 1204–1219. <https://doi.org/10.1111/1467-8624.00467>
- Rutter, M. (1987). Psychosocial resilience and protective mechanisms. *American Journal of Orthopsychiatry*, 57(3), 316–331. <https://doi.org/10.1111/j.1939-0025.1987.tb03541.x>
- Sabol, S. Z., Hu, S., & Hamer, D. (1998). A functional polymorphism in the monoamine oxidase A gene promoter. *Human Genetics*, 103(3), 273–279. <https://doi.org/10.1007/s004390050816>
- Sampson, R. J., & Laub, J. H. (1993). *Crime in the making: Pathways and turning points through life*. Harvard University Press.
- Scheff, T. J. (1988). Shame and conformity: The deference-emotion system. *American Sociological Review*, 53(3), 395–406. <https://doi.org/10.2307/2095647>
- Setoh, P., Zhao, S., Santos, R., Heyman, G. D., & Lee, K. (2020). Parenting by lying in childhood is associated with negative developmental outcomes in adulthood. *Journal of Experimental Child Psychology*, 189, Article 104680. <https://doi.org/10.1016/j.jecp.2019.104680>
- Shamay-Tsoory, S. G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain*, 132(3), 617–627. <https://doi.org/10.1093/brain/awn279>

- Shelton, K. K., Frick, P. J., & Wootton, J. (1996). Assessment of parenting practices in families of elementary school-age children. *Journal of Clinical Child Psychology*, 25(3), 317–329. https://doi.org/10.1207/s15374424jccp2503_8
- Shukla, M., & Upadhyay, N. (2025). Cold hearts and dark minds: A systematic review and meta-analysis of empathy across dark triad personalities. *Frontiers in Psychiatry*, 16, Article 1546917. <https://doi.org/10.3389/fpsy.2025.1546917>
- Shumay, E., Logan, J., Volkow, N. D., & Fowler, J. S. (2012). Evidence that the methylation state of the monoamine oxidase A (MAOA) gene predicts brain activity of MAO A enzyme in healthy men. *Epigenetics*, 7(10), 1151–1160. <https://doi.org/10.4161/epi.21976>
- Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology*, 24(18), R875–R878. <https://doi.org/10.1016/j.cub.2014.06.054>
- Summerhayes, D. (2026a). *Arriving at the centre of consciousness: A transjective framework* [Manuscript submitted for publication].
- Summerhayes, D. (2026b). *Empathy, compassion and the humanity paradox: A transjective account* [Preprint]. Zenodo. <https://doi.org/10.5281/zenodo.23600418>
- Summerhayes, D. (2026e). *MAOA-uVNTR, childhood adversity and differential susceptibility: Protocol for a systematic review and meta-analysis* [Manuscript in preparation].
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks/Cole.
- Thornberry, T. P., Krohn, M. D., Lizotte, A. J., Smith, C. A., & Tobin, K. (2003). *Gangs and delinquency in developmental perspective*. Cambridge University Press.
- van IJzendoorn, M. H. (1995). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. *Psychological Bulletin*, 117(3), 387–403. <https://doi.org/10.1037/0033-2909.117.3.387>
- Vigil, J. D. (1988). *Barrio gangs: Street life and identity in Southern California*. University of Texas Press.
- Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., Dymov, S., Szyf, M., & Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, 7(8), 847–854. <https://doi.org/10.1038/nn1276>
- Werner, E. E. (1993). Risk, resilience, and recovery: Perspectives from the Kauai Longitudinal Study. *Development and Psychopathology*, 5(4), 503–515. <https://doi.org/10.1017/S095457940000612X>
- Werner, E. E., & Smith, R. S. (1982). *Vulnerable but invincible: A longitudinal study of resilient children and youth*. McGraw-Hill.
- Widom, C. S. (1989). The cycle of violence. *Science*, 244(4901), 160–166. <https://doi.org/10.1126/science.2704995>