

Teaching Norms: Direct Evidence of Parental Transmission *

Thijs Brouwer[†] Fabio Galeotti[‡] Marie Claire Villeval[‡]

July 17, 2022

Abstract

We examine the educative role played by parents in social norm transmission. Using a field experiment, we study whether parents enforce and comply more with norms when their children are present compared to when they are not. We compare similar parents when or after they bring or pick up their children at school. We find that parents accompanying children, in contrast to parents alone, are more likely to punish norm violators and to provide help to strangers when there is no violation. They also tend to substitute more direct punishment with withholding help as a means of indirect punishment.

Keywords: Field Experiment, Social Norms, Transmission, Parenting, Norm Enforcement.

JEL codes: C93, D1, D63.

* *Acknowledgements:* We express our gratitude to the actors and research assistants for their excellent assistance in the project. We thank C. Bicchieri, S. Gaechter, E. Fatas, N. Jacquemet, P. Moffatt, H. Schildberg-Hörisch, R. Sugden, M. Sutter, B. Tungodden, and participants at the ANZWEE Conference in Melbourne, at the NOBEC Talks at the University of Pennsylvania, the Workshop on Norms and Image Concerns in Bonn, and seminar participants in Bonn, Dijon, Düsseldorf, Grenoble, Konstanz, Lund, Norwich, Paris, Reading, Tilburg, Vienna, and at WZB Berlin for helpful comments. Support from the LABEX CORTEX (ANR-11-LABX-0042) of Université de Lyon, within the program Investissements d'Avenir (ANR-11-IDEX-007) operated by Agence Nationale de la Recherche (ANR) and from IDEXLYON from Université de Lyon (INDEPETH-IDEX/SBP/2018/03) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005) is gratefully acknowledged. This project was conceived during T. Brouwer's research visit at GATE LSE, funded by a grant of IDEXLyon from the Université de Lyon.

[†]Tilburg University, Department of Economics. Email: thijs_brouwer@hotmail.com

[‡]Univ Lyon, CNRS, GATE, UMR 5824, F-69130 Ecully, France. Email: fabio.galeotti@cnrs.fr, villeval@gate.cnrs.fr

1 Introduction

Social norms, as shared understandings of what is acceptable or forbidden to do in society (Ostrom, 2000), play a major role in governing daily social interactions in various economically relevant settings. While there is an abundant literature on how they are enforced, we still know little on how they are transmitted from one generation to another (e.g., Bowles and Gintis, 1998; Lindbeck and Nyberg, 2006; Tabellini, 2008). Research in developmental psychology suggests that children acquire social behavior and internalize norms from young age partly through the observation and subsequent imitation of adults (Bandura, 1965, 1977). Parents are said to play a major role in this social learning process (McCord and McCord, 1958; Maccoby, 1992; Bauer et al., 2014). Norms can also be internalized by experiencing punishment when not complying with them (Sugden, 1986; Coleman, 1994; Young, 2008) or by observing others enforcing these norms (Malouff et al., 2009). Children understand the fairness of such enforcement (Piaget, 1932) and use third-party costly punishment already at primary school age (Lergetporer et al., 2014). This implies an alternative way through which norms are transmitted: parents may teach their child about norms by not only complying with them but also by *punishing* norm violations of others in the presence of their child. By doing so, the child learns vicariously that violations will not go unpunished, which may serve as an efficient deterrence mechanism to ensure future generation’s norm compliance.

We conducted a field experiment in the vicinity of 30 public elementary schools, chosen randomly, in Lyon, France, involving 601 parents of children aged 3 to 12 to address two questions in relation to parental transmission behavior. First, we ask whether parents engage more in norm enforcement in the presence of their child, presumably with the aim of educating the child. Our conjecture is that parents are willing to exert more effort in terms of norm enforcement in the presence of their child because it allows them to transmit to their offspring their values and principles about the right conduct, with the expectation of a higher benefit for their child in the future. This is related to the notions of parents’ imperfect empathy (Bisin and Verdier, 2001) and paternalistic altruism (Doepke and Zilibotti, 2019).

Second, we ask whether the nature of punishment changes in the presence of the child. Previous research has indeed identified two main forms of norm enforcement (Balafoutas et al., 2014). As a first means, someone can engage in direct punishment of a violation by verbally confronting the violator. However, the deterring effect of direct punishment is seriously hampered by the risk of retaliation (Janssen and Bushman, 2008; Nikiforakis, 2008), which typically causes direct punishment rates of norm violations to be low in field studies compared to anonymous laboratory experiments (Balafoutas and

Nikiforakis, 2012; Balafoutas et al., 2014, 2016; Berger and Hevenstone, 2016; Artavia-Mora et al., 2017). Instead, individuals may resort to indirect punishment that does not involve a risk of retaliation (Casari, 2012), as a substitute for direct punishment. In line with Balafoutas et al. (2014), we identify indirect punishment as withholding help to the violator. We explore whether the presence of their child changes the way parents use direct *vs.* indirect punishment of a norm violation to educate the child. We conjecture that the parents may not necessarily perceive that the risk of counter-punishment is higher in the presence of the child but that such counter-punishment may generate a negative externality by harming their image in the eyes of their child or by inducing some stress in the child. If this is the case, enforcing a norm directly is associated with a higher cost, which may lead to a higher degree of substitution between direct and indirect punishments.

Our primary focus is on the norm of non-littering and the violation thereof.¹ We employed a 3×2 design in which we varied the opportunity to enforce the norm and the presence of the child. To isolate the effect of the presence of a child, and thus the parents' teaching motivation, it is crucial to be able to compare similar parents in the same environment, the only difference being the presence of the child. We targeted parents who either brought their child to school in the morning or just picked up their child from school in the afternoon, and parents who either just left their child at school in the morning or came to pick up their child from school in the afternoon. This set-up provides natural variation in whether or not the child is present for an otherwise comparable sample of parents.

In order to implement the norm violation scenario in the most similar conditions across treatments and with high control, we recruited two trained actors, a male and a female. Each actor played one of three different scenes, inspired by Balafoutas et al. (2014). Depending on the scene, the actor (i) violated intentionally the non-littering norm ("Violation"), (ii) dropped, seemingly accidentally, the content of his or her bag on the ground, suggesting a need for spontaneous help ("Help"), or (iii) both ("Violation + Help"). These scenes provided the targeted parent an opportunity to sanction the norm violation directly, help a stranger in need, or both, respectively. Treatment (i) informs us on the prevalence of direct punishment, while treatments (ii) and (iii) together are informative of the rates of indirect punishment through withholding help.

We find that twice as many parents (22 *vs.* 11 percent) engaged in direct punish-

¹The importance of compliance with this norm is universally acknowledged. As an illustration, a representative survey ($N = 1060$) conducted by TNS Sofres revealed that 85% of French adult population considered littering unacceptable, and more deplorable than other acts such as vandalism (74%) and speeding on the highway (44%) (<https://www.recyclage-recuperation.fr/archives-dechets-com/les-francais-se-disent-anti-dechets-sauvages>; accessed on June 8, 2022)

ment when accompanying children, as compared to being alone. This reveals a higher willingness of parents to enforce the non-littering norm in the presence of their child. In the absence of a norm violation, parents with children were about twice more likely to help the actor pick up his or her fallen items compared to parents alone (47 *vs.* 26 percent), expressing the higher willingness of parents to model the norm of helping when their child is present. By contrast, the drop in helping following a violation was not significantly larger for parents accompanying children than for parents alone (16 *vs.* 10 percentage points). Our experiment does not provide conclusive evidence for a higher tendency to punish indirectly, although parents *were* more likely to substitute direct for indirect punishment when their child was around.

These findings contribute to the literature on the role of parents in the inter-generational transmission of values and preferences by assessing their norm enforcement behavior with the aim of educating their children. The development of preferences in children has received a lot of attention in the developmental psychology and economics literature (Sutter et al., 2019), resulting from the need to better understand the behavior of adults through the development of non-cognitive skills in childhood (e.g., Heckman and Rubinstein, 2001; Heckman et al., 2006). This has spurred a myriad of papers assessing the correlation between parents' and children's preferences (Dohmen et al., 2012; Kosse and Pfeiffer, 2012; Alan et al., 2017; Brenøe and Epper, 2019; Chowdhury et al., 2020; Cipriani et al., 2013; Sutter and Untertrifaller, 2020; Chowdhury et al., 2020). Factors that have been identified as affecting the formation and transmission of preferences include parenting styles and investments (Cunha and Heckman, 2010; Heckman and Mosso, 2014; Zumbuehl et al., 2013; Alan et al., 2017; Attanasio et al., 2020), breastfeeding as a proxy for the quality of early child environment (Falk and Kosse, 2016), exposure to a mentor (Kosse et al., 2020), language (Sutter et al., 2018), parental ambitions (Khadjavi and Nicklisch, 2018) and culture (Gneezy et al., 2009; Andersen et al., 2013; Falk et al., 2018). Furthermore, parental behavior has been identified as an important explanation for the gap in preferences and personality traits between children from low and high socio-economic status families (Benenson et al., 2007; Bauer et al., 2014; Deckers et al., 2017; Falk et al., 2018; Kosse et al., 2020; Sutter and Untertrifaller, 2020).

This literature shows the potential for children's preferences to be moulded by parental socialization efforts, but few scholars have examined this directly. Ben-Ner et al. (2017) investigated parents' tendency to model behavior in a dictator game to their children, Lindbeck and Nyberg (2006) and Houser et al. (2016) studied parents' dishonest behavior in the presence and absence of their child, and Cappelen et al. (2020) compared the effects of early childhood interventions and parenting programs on the

formation of social preferences in children. While these studies were conducted in highly controlled environments, we are able to assess parents’ natural behavior because parents and children were not aware of being part of an experiment. Moreover, we take a novel perspective by not just studying the extent to which parents model behavior, but also how they teach through punishing others’ undesired behavior. This latter feature has, to the best of our knowledge, not been studied before.²

Section 2 lays out our design and conjectures. Section 3 reports our results and section 4 provides a concluding discussion.

2 Experimental Design

We ran our field experiment in the vicinity of 30 randomly-selected public elementary schools in Lyon, France. We recruited two trained actors, one male and one female, from a professional acting school. Teams of four collected the data: one actor, two research assistants (RA1 and RA2), and one supervisor (one of the researchers). The actor and the two RAs were blind to the purpose of the study.³

2.1 Conditions and Conjectures

Our experiment uses a 3×2 between-subject design. One dimension varies the naturally occurring presence of the child, while the other dimension manipulates experimentally the setting of a norm violation. Regarding the first dimension, we targeted parents approaching *and* leaving the school in the morning *and* in the afternoon. This naturally varies whether the child is present or the parent is alone. This creates two conditions, to which we refer as “Child” (*C*) and “Alone” (*A*). This design makes sure that we target a similar sample of parents in the two conditions.

Regarding the second dimension, the protocol of the scenes is inspired by [Balafoutas et al. \(2014\)](#). In all conditions, the actor, who wears plain clothes, holds a small plastic bag containing food waste (a banana peel) and carries a cotton shoulder bag containing five file folders and a few pens (see [Appendix A.3](#)). The actor plays one of three

²We acknowledge that our study only documents the teaching side of the transmission process. Observing both teaching by parents and learning by their children would offer a complete picture of the transmission process duality. It would be extremely challenging, though, in a natural setting. Giving children an opportunity to violate a norm just after our intervention would have allowed us to test whether the children whose parents enforced the norm in reaction to the actor’s intervention reacted differently than children whose parents did not enforce it. However, such option would have raised ethical issues. An alternative would have been to survey the children about social norms. This would have raised other challenging issues, such as measuring the children’s initial knowledge of social norms and time pressure. Thus, we did not follow this line. Note, however, that studying parents’ teaching behavior is the first natural and indispensable step when analyzing the transmission process.

³Ethical approval was obtained from CEEI, the IRB of the French National Institute of Medical Research and Health (Inserm) (IRB00003888, No.19-592).

different scenes in front of a targeted parent in the streets surrounding the targeted school. The scenes constitute our three treatments and they are summarized in Table A1 in Appendix A.1. They differ in the opportunities that are provided to the parent to enforce the norm, as explained below. The first scene, called the “Violation” treatment, aims at measuring the prevalence of direct punishment of a social norm violation. In this scene, the actor approaches the targeted parent from the front. When the parent is roughly 10 meters away, the actor pauses and goes through the cotton bag. Then, when the parent is roughly 5 meters away, the actor litters in the clear sight of the parent by throwing away the plastic bag with the food waste towards the side of the street. Subsequently, the actor slowly starts moving again, while still going through the bag, clearly showing no intention to pick up the litter. We classify all forms of verbal confrontation aimed directly at the actor and which explicitly address the violation as direct punishment.⁴

The difference in direct punishment rates after observing the violation of the norm between the Child and Alone conditions informs on the parents’ tendency to engage in direct punishment with the goal to teach the child that the norm violation constitutes misbehavior that ought to be punished. Following social learning theory (Bandura, 1977) or parenting models such as Doepke and Zilibotti (2019), we conjecture that educative motives spur parents to inflict direct punishment when the child is present.⁵

Conjecture 1 (Direct Punishment). *In reaction to a social norm violation, direct punishment rates are higher for parents accompanying a child than for parents alone.*

An implicit assumption underlying Conjecture 1 is that the fear of retaliation is the same across conditions. This need not be the case. On the one hand, the parent may be afraid of retaliation more in the presence of the child. In particular, the parent may fear for his or her image and be anxious that the child will be involved in some way in the retaliation, even by simply witnessing it. If this is the case, our experiment may underestimate the effect of the child’s presence and this would work against Conjecture 1, making any results that support it even more convincing. On the other hand, the parent may deem the violator less likely to retaliate in front of a child. If so, we may overestimate the effect of the child’s presence. We discuss this issue further in Section 3 when presenting our results.

Because of the fear of retaliation discussed above, direct punishment rates are typically low in the field (Balafoutas et al., 2014, 2016). Instead of direct punishment, parents may resort to forms of indirect punishment, for which retaliation is arguably less likely. We provide the opportunity for indirect punishment in the second treatment,

⁴For example, “Do not throw that on the ground; throw it in the garbage bin.”

⁵All conjectures have been pre-registered with AsPredicted (#24270).

called “Violation + Help”. This scene starts in a similar way as the Violation scene. Then, after having littered, the actor accidentally drops the contents of his or her bag on the sidewalk. This presents the parent with an opportunity to withhold help as a form of indirect punishment.⁶

We define a parent to help if he or she picks up at least one item from the ground, as in Balafoutas et al. (2014). In case parents stimulate their children to help, this also counts as helping.⁷ In order to measure whether parents withhold help, we introduce a third treatment, called “Help”. In this scene, the helping opportunity is not preceded by a littering violation. Indirect punishment then shows in the aggregate through significantly lower helping rates in the presence of a violation. We conjecture that educative motives induce parents to punish indirectly more often when the child is around, meaning that we should observe a larger decrease in helping rates in the presence, rather than the absence, of the child.

Conjecture 2 (Indirect Punishment). *The extent to which parents withhold help after observing the violation of a social norm is larger in the presence of the child.*

Although indirect punishment is less likely to evoke retaliation, the educative motive of withholding help is probably weaker than that of direct punishment, as its implicit nature may be harder for the child to grasp. The child may thus not understand completely that the parent is punishing the violator, especially if the helping norm is not yet well-established in the child’s mind. Hence, the marginal teaching benefit from indirect punishment is likely to be smaller than that of direct punishment. This consideration would, however, work against our conjecture. The Violation + Help treatment may also be informative about the fear of retaliation between conditions, as we can observe the extent to which direct punishment is substituted for indirect punishment. In case the drop in direct punishment is larger (smaller) in the Child condition as compared to the Alone condition, this would point to a higher (lower, respectively) fear of retaliation in the presence of the child.

Finally, helping someone is another social norm that parents may be willing to transmit to their children. Based on social learning theory (Bandura, 1977), we expect helping rates in the Help scene to be higher for parents accompanying their child, as compared to parents alone.

⁶To equalize the window of opportunity to punish *directly* across scenes, the actor pauses a second time in the Violation scene around the same time where (s)he would drop the files in the other two scenes.

⁷This was a rare event and treating it as not helping does not change the results (see Appendix A.4.5). The rare cases the child helped without any intervention of the parent are not counted as helping.

Conjecture 3 (Helping). *When no social norm is violated, parents in the presence of their child, in contrast to parents alone, are more likely to provide help.*

2.2 Procedures

The different scenes were played in mixed order in the morning approximately between 7:45 and 9:00, and in the afternoon between 16:15 and 18:00. We visited each school during at most one day to avoid being recognized. The actors were randomly alternated across schools, except for the first session where both actors were present in order to homogenize the procedures.⁸ Targeted parents were identified by the actor and the supervisor. We restricted ourselves to single parents walking alone or with one or more children. To make the cost of helping negligibly small for parents and unrelated to the presence of the child, and because being in emergency might affect helping behavior (Darley and Batson, 1973), we did not target parents visibly in a rush, talking on the phone, pushing a stroller, riding a bike, walking a dog, holding something with both hands, and/or accompanying disabled children. In case the parent engaged in direct punishment, the actor always complied and disposed of the litter. Subsequently, the actor quietly left the scene.

In the meantime, RA1 observed the scene from a distance and recorded the type of scene being played, and whether the target actually observed the actor littering and dropping the content of his or her bag, because it was crucial that targets paid attention to the scene being played. RA1 also recorded whether the parent was accompanied by one or more children, and the parent's response to the scene: whether the targeted parent helped the actor, confronted verbally the actor regarding the violation, or expressed his or her disapproval to the child in a way that could be heard by the actor. RA1 recorded whether there were witnesses who could possibly have intervened in the scene, the gender of the parent and the child, the weather conditions, the time of the day, and the cleanliness of the street. After the scene had ended, RA1 cleaned up the scene in case of a non-sanctioned littering violation and verified the recorded information with the actor.

After the targeted parent had left the scene, RA2 approached the parent and asked whether he or she was willing to participate in a short, seemingly unrelated, survey. Our main interest was in the question whether the target was accompanying or had accompanied (i) his or her own child, (ii) a child that he or she guards, (iii) the child of a relative, or (iv) no child. In cases (i) to (iii), we also asked for the child(ren)'s age and gender. If a parent declined to take the survey, RA2 had to guess the gender and

⁸In three occasions, one actor could only be present in the morning, and was substituted by the other actor in the afternoon.

age of the child(ren), if present. We dropped from the sample targets indicating they had no child going to the school. For an exhaustive list of information recorded, see Appendix A.2.

Challenges We identified three main challenges in relation to our design. First, parents might be more in a hurry when arriving at school, rather than when leaving school. We addressed this issue by running the scenes both in the morning when parents dropped off their kids *and* in the afternoon when parents arrived at school to pick up their child. As a result, we equalized parents’ “hurry” between child conditions as much as possible. As additional measures to avoid parents in a hurry altogether, we mainly staged the scenes before the five minutes prior to the beginning or end of the class, and avoided parents who were visibly in a rush. Also, we ensured that each of the scenes did not last for more than a couple of seconds and that helping constituted a quick act. Note that direct punishment does not take long either since the target can express disapproval while continuing to walk. Finally, we examined whether there are timing effects in our data and found none (see Appendix A.4.7).

A second concern was the audience of the interaction. It was hard to target only one parent in an otherwise empty street, because parents arrived with their children around the same time. In addition to RA1 recording whether there are any witnesses, we combated this by identifying the more secluded streets in the school neighborhoods and be present 30 to 40 minutes before the beginning or ending of class. In case another parent was approaching while the scene was about to start, the actor waited until the person had passed before starting the scene.

A third concern was whether the targeted adult was indeed the parent of the child. Children might also be picked up from and dropped off at school by their nannies or other caretakers. Insofar as these caretakers still play an important role in the child’s education, studying their behavior remains relevant. We might, however, need to be careful with labeling observed behavior as a tendency of *parents per se*. We combated this potential issue by means of the survey conducted directly after the scene and that allowed us to evaluate the proportion of parents across conditions.

Locations and Sample Size In Appendix A.1, we discuss the determination of the sample size and the selection of the schools in detail.

3 Results

3.1 Data Description

The experiment was run on 30 days in May and June 2019. We collected 601 observations: 301 in the Alone condition and 300 in the Child condition. Our primary outcome variables are (Direct) Punishment and Helping. Punishment is a dummy that takes on value 1 if the parent verbally and explicitly punishes the actor for the violation of the non-littering norm, and 0 otherwise. Helping is a dummy that takes on value 1 if the parent picks up at least one item (or asks the child to pick one), and zero otherwise.

Our main interest is in the effect of the scene played (Violation, Help, or Violation + Help) and the condition (Child or Alone). Moreover, we control for the following observables in the regressions. Male Target is a dummy indicating that a father was targeted, while Male Actor indicates that the male actor played the scene. Morning, Rain, and Hot are dummies indicating that the scene was played during the morning, during rainy conditions, and on a hot day, respectively. Witness is a dummy that takes on value 1 if RA1 deemed another non-targeted adult to be observing the scene *and* to be able to intervene. Finally, we created Rich IRIS which takes value 1 if the school’s IRIS area median income is above the city-wide median.⁹ Table 1 shows that our sample is balanced on most controls, with two notable exceptions. The scenes of the Child condition were somewhat more likely to be played in the morning (significant at the 1% level) and with a witness around (significant at the 10% level).

Regarding the survey, 47% of the approached parents responded to the survey. Here, we exclude parents who could not be reached for various reasons (*e.g.*, they were talking to another person). Responding to the survey can to some extent be treated as a helping act. Since parents in the Child condition were more likely to respond to the survey than parents in the Alone condition, this may hint at parents behaving more pro-socially in the presence of their child. It may also indicate that parents in the Alone condition were on average more in a hurry than when accompanied; however, there is no obvious reason for which this would be the case, and the higher response rate of parents with their child was observed regardless of the time of the day and of the condition (see Appendix section A.4.1). Nevertheless, the data analysis accounts for the parents’ time of arrival (see Challenge 1). Moreover, as shown in Appendix section A.4.1, the response rates varied across treatments: a higher proportion of parents responded after observing the actor needing help than after observing a norm violation, which suggests that there might be some spillover effects of the scene on the survey response rate.¹⁰ Importantly,

⁹IRIS are infra-municipal areas comprising between 1800 and 5000 residents.

¹⁰For a study on spillover effects across contexts see Galeotti et al. (2021)

of those who answered, the vast majority (88%) reported being the parent of the child, rather than a guardian, with no difference between conditions (see Table A3 in Appendix A.4.1).

Table 1: Summary Statistics

	(1) ALL	(2) ALONE	(3) CHILD	(4) DIFF.
Male Target	0.35 (0.48)	0.36 (0.48)	0.34 (0.47)	0.02 (0.04)
Witness	0.15 (0.35)	0.12 (0.33)	0.17 (0.38)	-0.05* (0.03)
Male Actor	0.51 (0.50)	0.52 (0.50)	0.50 (0.50)	0.02 (0.04)
Rich IRIS	0.51 (0.50)	0.50 (0.50)	0.51 (0.50)	-0.01 (0.04)
Morning	0.52 (0.50)	0.46 (0.50)	0.58 (0.49)	-0.13*** (0.04)
Rain	0.08 (0.27)	0.09 (0.29)	0.07 (0.25)	0.02 (0.02)
Hot	0.28 (0.45)	0.31 (0.46)	0.26 (0.44)	0.05 (0.04)
Survey Response ^a	0.47 (0.50)	0.42 (0.49)	0.51 (0.50)	-0.09** (0.04)
Observations	601	301	300	601

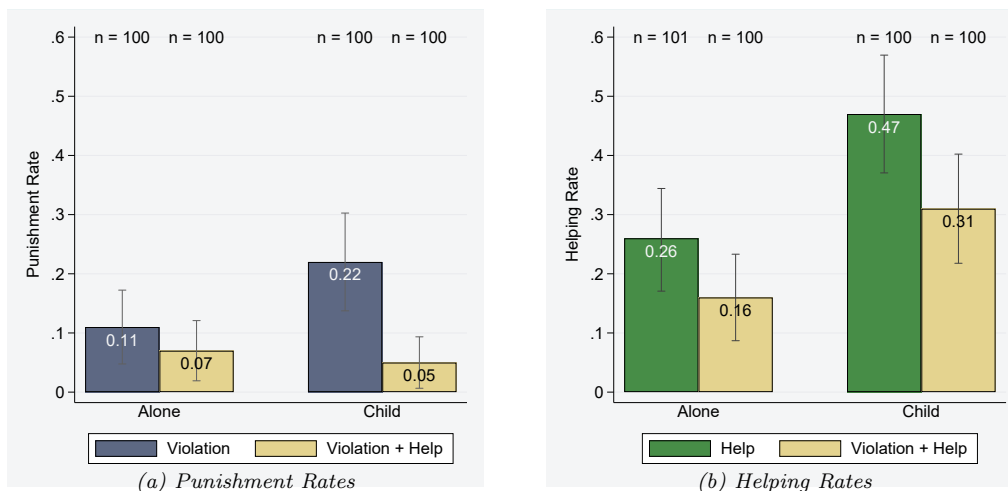
Notes: Columns (1)-(3) contain standard deviations in parentheses. Column (4) contains standard errors in parentheses. *a*: Parents who could not be reached are excluded. Hence, the statistics are computed based on 503, 251, 252 observations in All, Alone, and Child, respectively. All tests are two-sided t-tests on the equality of means. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

3.2 Main Results

The two panels of Figure 1 display parents' direct punishment rate and helping rate, respectively, by treatment (Violation or Help, and Violation + Help) and by condition (Alone or with Child). To test our conjectures and compare the conditions we employ χ^2 -tests. Moreover, Table 2 reports the estimates of two linear probability models for each of the dependent variables, Punishment and Helping. Models (1) and (3) estimate a simple model including a dummy for the Child condition, a dummy for the Violation + Help scene (VH), and an interaction term. To allow for observations at the same school to be correlated, standard errors are clustered at the school-level. Models (2) and (4) also include the variables in Table 1 (except Survey Response) as controls. We ignore the Help (Violation) scene when analyzing Punishment (Helping, respectively), thus analyzing roughly 400 observations.

The direct punishment rate in the Violation treatment is 22 percent in the Child condition and 11 percent in the Alone condition (compare the dark bars in Figure

Figure 1: Behavior of Parents, by Treatment and Condition



1a). We reject the null of no differences in punishment rates between conditions ($\chi_1^2 = 4.39, p = 0.036$), showing evidence in line with Conjecture 1. This result is backed up by the linear probability model, as shown by the positive and significant coefficient on Child in the left panel of Table 2. This analysis supports our first result:

Result 1 (Direct Punishment). *Parents accompanying children are significantly more likely to engage in direct punishment following the violation of a social norm.*

This result is consistent with the willingness of parents to punish more when the child is around in order to teach the child about the importance of norm compliance and the risk of being sanctioned in case of a violation. We can reject four alternative explanations.

To begin with, we are not estimating a social image effect (*i.e.*, parents punishing or helping purely because they know that they are being observed). Indeed, the coefficient of the Witness dummy in Table 2 (that captures the pure presence of another adult observing) is negative and insignificant in both model (2) and (4): being observed by another adult does not increase the parents' likelihood of punishing the violator or their willingness to help. If we include an interaction term between the Child and Witness variables, this does not change the sign of the coefficient of the Witness dummy (we report this analysis in Appendix A.4.2). In addition, the interaction term is never significant: parents do not enforce more the norm when the child is present because they are observed by another adult. We also estimated models similar to those reported in Table 2 but treating the presence of a witness as if it was a treatment variable in the Alone condition. Consistently with the former analysis, we found that the Witness variable has a significant and negative effect on the likelihood of punishing and a negative but non-significant effect on the likelihood of helping (whereas the Child variable in Table 2 is significant and positive in both models). The 95% confidence intervals

Table 2: Determinants of Punishment Rate (Left) and Helping Rate (Right)

	PUNISHMENT		HELPING	
	(1)	(2)	(3)	(4)
Child	0.11** (0.05)	0.12** (0.04)	0.21*** (0.05)	0.22*** (0.06)
VH	-0.04 (0.03)	-0.04 (0.03)	-0.10** (0.05)	-0.10** (0.04)
Child × VH	-0.13*** (0.04)	-0.12*** (0.04)	-0.06 (0.07)	-0.07 (0.08)
Male Target		0.06** (0.03)		0.03 (0.04)
Male Actor		-0.06 (0.04)		-0.22*** (0.05)
Morning		-0.04 (0.03)		0.01 (0.04)
Witness		-0.04 (0.04)		-0.13 (0.08)
Rich area		0.02 (0.04)		0.01 (0.05)
Rain		0.01 (0.04)		-0.08 (0.07)
Hot		-0.04 (0.05)		0.06 (0.05)
Constant	0.11*** (0.03)	0.14*** (0.05)	0.26*** (0.04)	0.36*** (0.05)
Observations	400	399	401	400
Clusters	30	30	30	30
R^2	0.04	0.07	0.06	0.13
F	7.74	5.17	7.36	6.45
df	29	29	29	29

Notes: The table contains results from pooled Ordinary Least Squared regressions. The dependent variable is a binary variable for punishment (columns 1 and 2) or helping (columns 3 and 4). Standard errors in parentheses are clustered at the school level (30 clusters). One observation is dropped due to missing data on the target’s gender. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

are $[-0.198, -0.012]$ and $[-0.475, 0.058]$ for Punishment and Helping, respectively. For comparison, for the Child dummy in the original regressions, these are $[0.105, 0.338]$ and $[0.026, 0.206]$ for Helping and Punishment, respectively. Overall, this analysis reveals that the presence of another adult has either no effect or the opposite effect compared to the presence of a child: individuals feel less committed to punish norm violations when there are witnesses (in line with the “bystander effect”).

Second, we can also reject that higher punishment in the presence of the child is driven by the parents’ willingness to enhance their image in the eyes of their own child. The drop in direct punishment rates when comparing the Violation and Violation + Help treatments in the Child condition, as shown in Figure 1a and Table 2, invalidates this argument. Indeed, if parents punished more to show their child that they are strong,

this should remain the case in the Violation + Help treatment.

Third, we can reject an explanation in terms of lower fear of retaliation in the presence of the child (*i.e.*, parents being more likely to punish because they believe that a violator is less inclined to retaliate in front of a child). Indeed, punishment decreases significantly and substantially between the Violation and Violation + Help treatments in the Child condition *only*, as can be seen in Figure 1a (Child: $\chi_1^2 = 12.37, p < 0.001$; Alone: $\chi_1^2 = 0.98, p = 0.323$). This shows that only parents with children decrease direct punishment when indirect punishment opportunities are available. This suggests that parents with children fear retaliation *more*.¹¹

Finally, we are able to reject the alternative explanation that the perception of the social norm changes with the presence of the child. For example, parents may deem the violation a more serious transgression in the presence of the child, which might increase their tendency to punish. In order to assess this, we first conducted an online vignette study in Lyon with a different sample of parents to elicit social norms in the presence and absence of the child in scenarios close to those used in our experiment (using the design by Krupka and Weber, 2013, not pre-registered). We found no evidence that the social norm is different in the two conditions, as developed in Appendix A.5. Then, we conducted a second online study in which we presented to different participants one of three recorded videos featuring a person littering in a public space in front of a bystander. In one video the bystander was alone, while in the two other videos the bystander was accompanied, either by a child or by another adult. To elicit the injunctive social norms and personal norms, the participants were asked to guess the modal ratings of the littering violation and the subsequent verbal punishment by the bystander (not shown in the video). This study and its results are presented in detail in Appendix A.6. The second study confirms that parents in our field study did not punish more in the presence of a child because they considered littering a more serious violation or because they deemed direct punishment more appropriate in the presence of the child. Since we found again no significant differences in the norms across conditions, this means that individuals do not deem the littering norm violation more serious in the presence of a child than without.¹² Taken together, we thus adopt the teaching motive as the dominant explanation of our results.

¹¹An additional analysis of the Violation treatment shows that only fathers punish significantly more in the presence of the child (see Table A5 in Appendix A.4.3). Assuming that fathers fear retaliation less overall, the presence of the child should decrease the gap in punishment rates between mothers and fathers if children would reduce the fear of retaliation. If anything, we find the opposite.

¹²Note that if we had found significant differences, it would have suggested that the stronger norm enforcement in the presence of the child in our field experiment might have been driven by the parents' willingness to teach the violator that littering is more inappropriate in the presence of a child. This would still have pointed on norm enforcement as a transmission mechanism, but the target might have been more the violator than the child.

Before we move to the analysis of indirect punishment, we note that helping rates in the Help treatments are significantly higher in the Child condition, as compared to the Alone condition (compare the dark bars in Figure 1b): 47 *vs.* 26 percent ($\chi_1^2 = 9.82, p = 0.002$). This is confirmed by the positive and significant coefficients on Child in models (3) and (4) of Table 2, and it is consistent with the willingness to teach children compliance with the helping norm. Result 2 is thus in line with Conjecture 3.

Result 2 (Helping). *Parents accompanying children are significantly more likely to provide help to a stranger in need.*

In both conditions, parents decrease their willingness to help following a norm violation. The helping rate decreases by 16 percentage points in the Child condition and by 10 percentage points in the Alone condition. Still, this rate remains significantly higher in the Violation + Help treatment in the Child condition: 31 *vs.* 16 percent ($\chi_1^2 = 6.26, p = 0.012$). Note that the helping rate in the Alone condition is similar to Balafoutas et al. (2014) and Balafoutas et al. (2016), where it was 18.6% and 13.0%, respectively. In order to test Conjecture 2 formally, we examine the interaction term $VH \times Child$ in our regressions. In models (3) and (4) of Table 2 the coefficient estimates on $VH \times Child$ suggest that the helping rate does not decrease *more* in the Child condition as compared to the Alone condition, even though the negative coefficient estimate is in the predicted direction. In sum, we find no statistical evidence for Conjecture 2. This leads to Result 3.

Result 3 (Indirect Punishment). *Parents accompanying children are not significantly more likely to engage in indirect punishment.*

From Figure 1a, it becomes apparent that parents withhold help as a substitute for direct punishment, as exemplified by the drop in punishment once a helping opportunity is presented, but substantially more so in the Child condition. The drop in direct punishment from 22 percent in Violation to 5 percent in Violation + Help is significant at the 1% level for the Child condition ($\chi_1^2 = 12.37, p < 0.001$). The drop from 11 to 7 percent in the Alone condition is insignificant ($\chi_1^2 = 0.98, p = 0.323$). As mentioned earlier, this suggests that parents may fear retaliation more when they are with their child, as they seem more eager to resort to indirect, rather than direct, punishment. Hence, because of this likely higher fear of retaliation, our result on direct punishment may underestimate parents' true tendency to punish more in the presence of the child.

Our results are robust to using logit models instead of linear probability models (see Appendix A.4.4) and to a number of sample restrictions and alternative definitions, as summarized in Table A7 in Appendix A.4.5. Furthermore, we performed some

exploratory (not pre-registered) regressions on heterogeneous treatment effects in Appendix A.4.6. We examine the effect of the number of children and their gender and age, the gender of the actor and parent, and the relative income in the school’s IRIS area on helping and punishment rates. Interestingly, we find that parents punish more in the presence of children aged 6 to 8, while they help more in the presence of children aged below 6. This suggests that parents seek to first establish the helping norm, which may be easier to understand at a young age (Warneken and Tomasello, 2006), and teach through punishment at a later age. This observation also contributes to reject hurry as a reason underlying the differences in behavior across conditions because hurry is likely independent of the child’s age. Furthermore, we show that the timing of the scene does not affect Punishment or Helping rates, although they should if our results were driven by parents who are in a rush, and that there is no significant effect of the interaction between the time slot and the presence of the child (see Appendix A.4.7).

4 Discussion and Conclusion

Our field experiment provides evidence of parental involvement in childhood education on social norm compliance through the enforcement of such norms, and thereby contributes to our understanding of the intergenerational transmission of normative preferences from parents to children. We find that parents punish a norm violation directly more often in the presence of their child, while they do not punish more often through withholding help. Withholding help, even from a norm violator, violates another norm: helping others. Thus, parents may believe that its educative signal is unlikely to be grasped by children and prioritize teaching the helping norm. This may explain why the helping rate remains relatively high after a violation in presence of the child.

By focusing on normative preferences and by assessing parents’ *natural* enforcement behavior, we complement studies showing the importance of the cultural transmission of preferences (Tabellini, 2008; Bisin and Verdier, 2001) and how preferences evolve during childhood. In line with the recent economic studies on parenting (Doepke and Zilibotti, 2017, 2019) and parental socialization efforts (Ben-Ner et al., 2017; Houser et al., 2016; Cappelen et al., 2020; Sutter and Untertrifaller, 2020), we show that parents exhibit more socially responsible behavior in the presence of their child. A novel aspect of our study is showing that parents not only model desired behavior, but also teach through punishing undesired behavior by other parties. This illustrates that parents seek to build their children’s social skills not only through direct interventions on their behavior but also by acting on their environment.

As explained above, we are able to reject alternative explanations related to the perception of the social norm, image concerns, or a lower fear of retaliation. However, admittedly, parents may be more sensitive to their environment or moving more slowly when with their child(ren). Even though the window of opportunity may be slightly smaller for parents walking alone, our design ensures that the parent always crossed paths with the actor after (s)he littered or dropped the bag. We excluded observations with parents who did not observe the violation, while the dropping of the items was difficult to miss. Moreover, direct punishment does not take time, as expressing verbally one’s disagreement does not stop the parents in their walk. It could still be that parents have a different mindset when with their children, but we argue that this is part of our teaching explanation. It might be interesting to compare how these adults would behave if accompanied by another adult or a non-kin child instead of their child. A recent study (Fornwagner and Hauser, 2020) found, however, that individuals’ contributions to a voluntary climate action was mainly sensitive to observability by their own child, compared to a non-kin child or another adult.

Incidentally, we note that the teaching motive may change the nature of punishment of violations: such punishment may no longer be completely altruistic if both current benefits to oneself (transmission of values to one’s offspring) and future benefits to the child (accumulated social skills) are expected. Our results show that these expected benefits compensate the cost associated with the threat of retaliation. A higher fear of retaliation in the presence of the child is supported by the tendency of accompanied parents to substitute more direct for indirect punishment, which also suggests that our Violation treatment may underestimate the teaching motivation of parents in our study.

Our study also reveals the heterogeneity of preferences across parents. Not all parents use the observed norm violation as an opportunity to teach the importance of norm compliance to their children. Although we did not find differences across districts in the city, our study calls for further investigations of the heterogeneity in the inter-generational transmission of norms and parenting styles, by examining the individual and institutional determinants of the degree of involvement of parents in teaching normative preferences to their children.

Our study suggests that heterogeneity in teaching, and not necessarily only in parents’ preferences, may play a role in producing diversity in the formation of normative preferences during childhood. It might be interesting to study whether parents’ behavior in our settings is correlated with what has been described as an authoritarian or as a permissive parenting style (Doepke and Zilibotti, 2019). Other interesting extensions would be to connect teaching and learning, to test whether children who have just been

taught through example are more likely to exhibit stronger normative preferences, and whether parents and other adults' teaching make a difference in such endeavor.

References

- ALAN, S., N. BAYDAR, T. BONEVA, T. F. CROSSLEY, AND S. ERTAC (2017): "Transmission of risk preferences from mothers to daughters," *Journal of Economic Behavior & Organization*, 134, 60–77.
- ANDERSEN, S., S. ERTAC, U. GNEEZY, J. A. LIST, AND S. MAXIMIANO (2013): "Gender, competitiveness, and socialization at a young age: evidence from a matrilineal and a patriarchal society," *Review of Economics and Statistics*, 95, 1438–1443.
- ANGERER, S., D. GLÄTZLE-RÜTZLER, P. LERGETPORER, AND M. SUTTER (2015): "Donations, risk attitudes and time preferences: a study on altruism in primary school children," *Journal of Economic Behavior & Organization*, 115, 67–74.
- ARTAVIA-MORA, L., A. S. BEDI, AND M. RIEGER (2017): "Intuitive help and punishment in the field," *European Economic Review*, 92, 133–145.
- ATTANASIO, O., S. CATTAN, E. FITZSIMONS, C. MEGHIR, AND M. RUBIO-CODINA (2020): "Estimating the production function for human capital: results from a randomized controlled trial in Colombia," *American Economic Review*, 110, 48–85.
- BALAFOUTAS, L. AND N. NIKIFORAKIS (2012): "Norm enforcement in the city: a natural field experiment," *European Economic Review*, 56, 1773–1785.
- BALAFOUTAS, L., N. NIKIFORAKIS, AND B. ROCKENBACH (2014): "Direct and indirect punishment among strangers in the field," *Proceedings of the National Academy of Sciences*, 111, 15924–15927.
- (2016): "Altruistic punishment does not increase with the severity of norm violations in the field," *Nature Communications*, 7, 1–6.
- BANDURA, A. (1965): "Vicarious processes: a case of no-trial learning," in *Advances in Experimental Social Psychology*, Elsevier, vol. 2, 1–55.
- (1977): *Social learning theory*, Oxford: Prentice Hall.
- BAUER, M., J. CHYTILOVÁ, AND B. PERTOLD-GEBICKA (2014): "Parental background and other-regarding preferences in children," *Experimental Economics*, 17, 24–46.
- BEGGS, S., S. CARDELL, AND J. HAUSMAN (1981): "Assessing the potential demand for electric cars," *Journal of econometrics*, 17, 1–19.
- BEN-NER, A., J. A. LIST, L. PUTTERMAN, AND A. SAMEK (2017): "Learned generosity? An artefactual field experiment with parents and their children," *Journal of Economic Behavior & Organization*, 143, 28–44.
- BENENSON, J. F., J. PASCOE, AND N. RADMORE (2007): "Children's altruistic behavior in the dictator game," *Evolution and Human Behavior*, 28, 168–175.

- BERGER, J. AND D. HEVENSTONE (2016): “Norm enforcement in the city revisited: an international field experiment of altruistic punishment, norm maintenance, and broken windows,” *Rationality and Society*, 28, 299–319.
- BISIN, A. AND T. VERDIER (2001): “The economics of cultural transmission and the dynamics of preferences,” *Journal of Economic Theory*, 97, 298–319.
- BOWLES, S. AND H. GINTIS (1998): “The Moral Economy as Community: Structured Populations and the Evolution of Prosocial Norms,” *Evolution and Human Behavior*, 19, 3–25.
- BRENØE, A. A. AND T. EPPER (2019): “Parenting values moderate the intergenerational transmission of time preferences,” *University of Zurich, Department of Economics, Working Paper n. 333*.
- CAPPELEN, A. W., J. A. LIST, A. SAMEK, AND B. TUNGODDEN (2020): “The effect of early education on social preferences,” *Journal of Political Economy*, in press.
- CASARI, M. (2012): “Weak reciprocity alone cannot explain peer punishment,” *Behavioral and Brain Sciences*, 35, 21–22.
- CHOWDHURY, S., M. SUTTER, AND K. F. ZIMMERMANN (2020): “Economic preferences across generations and family clusters: A large-scale experiment,” Tech. rep., GLO Discussion Paper.
- CIPRIANI, M., P. GIULIANO, AND O. JEANNE (2013): “Like mother like son? Experimental evidence on the transmission of values from parents to children,” *Journal of Economic Behavior & Organization*, 90, 100–111.
- COLEMAN, J. S. (1994): *Foundations of social theory*, Cambridge MA: Harvard University Press.
- CUNHA, F. AND J. J. HECKMAN (2010): “Investing in our young people,” *NBER Working Paper n. 16201*.
- DARLEY, J. M. AND C. D. BATSON (1973): “From Jerusalem to Jericho: A study of situational and dispositional variables in helping behavior,” *Journal of Personality and Social Psychology*, 27, 100–108.
- DECKERS, T., A. FALK, F. KOSSE, P. R. PINGER, AND H. SCHILDBERG-HÖRISCH (2017): “Socio-economic status and inequalities in children’s IQ and economic preferences,” *IZA Discussion Paper n. 11158*.
- DOEPKE, M. AND F. ZILIBOTTI (2017): “Parenting with style: altruism and paternalism in intergenerational preference transmission,” *Econometrica*, 85, 1331–1371.
- (2019): *Love, Money and Parenting*, Princeton University Press.
- DOHMEN, T., A. FALK, D. HUFFMAN, AND U. SUNDE (2012): “The intergenerational transmission of risk and trust attitude,” *Review of Economic Studies*, 79, 645–677.
- FALK, A., A. BECKER, T. DOHMEN, B. ENKE, D. HUFFMAN, AND U. SUNDE (2018): “Global evidence on economic preferences,” *The Quarterly Journal of Economics*, 133, 1645–1692.

- FALK, A. AND F. KOSSE (2016): “Early childhood environment, breastfeeding and the formation of preferences,” *SOEPpaper n. 882*.
- FORNWAGNER, H. AND O. P. HAUSER (2020): “Do parents invest into voluntary climate action when their children are watching? Evidence from a lab-in-the-field experiment,” *Working Papers in Economics and Statistics*, 2020-23.
- GALEOTTI, F., V. MAGGIAN, AND M. C. VILLEVAL (2021): “Fraud Deterrence Institutions Reduce Intrinsic Honesty,” *The Economic Journal*, 131, 2508–2528.
- GNEEZY, U., K. L. LEONARD, AND J. A. LIST (2009): “Gender differences in competition: evidence from a matrilineal and a patriarchal society,” *Econometrica*, 77, 1637–1664.
- HAUSMAN, J. A. AND P. A. RUUD (1987): “Specifying and testing econometric models for rank-ordered data,” *Journal of econometrics*, 34, 83–104.
- HECKMAN, J. AND Y. RUBINSTEIN (2001): “The importance of noncognitive skills: Lessons from the GED testing program,” *American Economic Review*, 91, 145–149.
- HECKMAN, J., J. STIXRUD, AND S. URZUA (2006): “The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior,” *Journal of Labor Economics*, 24, 411–482.
- HECKMAN, J. J. AND S. MOSSO (2014): “The economics of human development and social mobility,” *Annual Review of Economics*, 6, 689–733.
- HOUSER, D., J. A. LIST, M. PIOVESAN, A. SAMEK, AND J. WINTER (2016): “Dishonesty: from parents to children,” *European Economic Review*, 82, 242–254.
- JANSSEN, M. A. AND C. BUSHMAN (2008): “Evolution of cooperation and altruistic punishment when retaliation is possible,” *Journal of Theoretical Biology*, 254, 541–545.
- KHADJAVI, M. AND A. NICKLISCH (2018): “Parents’ ambitions and children’s competitiveness,” *Journal of Economic Psychology*, 67, 87–102.
- KOSSE, F., T. DECKERS, P. PINGER, H. SCHILDBERG-HÖRISCH, AND A. FALK (2020): “The formation of prosociality: causal evidence on the role of social environment,” *Journal of Political Economy*, 128, 434–467.
- KOSSE, F. AND F. PFEIFFER (2012): “Impatience among preschool children and their mothers,” *Economics Letters*, 115, 493–495.
- KRUPKA, E. L. AND R. A. WEBER (2013): “Identifying social norms using coordination games: why does dictator game sharing vary?” *Journal of the European Economic Association*, 11, 495–524.
- LERGETPORER, P., S. ANGERER, D. GLÄTZLE-RÜTZLER, AND M. SUTTER (2014): “Third-party punishment increases cooperation in children through (misaligned) expectations and conditional cooperation,” *Proceedings of the National Academy of Sciences*, 111, 6916–6921.

- LINDBECK, A. AND S. NYBERG (2006): “Raising children to work hard: altruism, work norms, and social insurance,” *The Quarterly Journal of Economics*, 121, 1473–1503.
- MACCOBY, E. E. (1992): “The role of parents in the socialization of children: an historical overview,” *Developmental Psychology*, 28, 1006–1017.
- MALOUFF, J., E. THORSTEINSSON, N. SCHUTTE, AND S. E. ROOKE (2009): “Effects of vicarious punishment: A meta-analysis,” *The Journal of general psychology*, 136, 271–286.
- MCCORD, J. AND W. MCCORD (1958): “The effects of parental role model on criminality,” *Journal of Social Issues*, 14, 66–75.
- NIKIFORAKIS, N. (2008): “Punishment and counter-punishment in public good games: can we really govern ourselves?” *Journal of Public Economics*, 92, 91–112.
- OSTROM, E. (2000): “Collective Action and the Evolution of Social Norms,” *The Journal of Economic Perspectives*, 14, 137–158.
- PIAGET, J. (1932): *The moral judgement of the child*, Routledge and Kegan Paul.
- SUGDEN, R. (1986): *The economics of rights, cooperation and welfare*, Oxford: Basil Blackwell.
- SUTTER, M., S. ANGERER, D. GLÄTZLE-RÜTZLER, AND P. LERGETPORER (2018): “Language group differences in time preferences: evidence from primary school children in a bilingual city,” *European Economic Review*, 106, 21–34.
- SUTTER, M. AND A. UNTERTRIFALLER (2020): “Children’s heterogeneity in cooperation and parental background: an experimental study,” *Journal of Economic Behavior & Organization*, 171, 286–296.
- SUTTER, M., C. ZOLLER, AND D. GLÄTZLE-RÜTZLER (2019): “Economic behavior of children and adolescents – a first survey of experimental economics results,” *European Economic Review*, 111, 98–121.
- TABELLINI, G. (2008): “The scope of cooperation: values and incentives,” *The Quarterly Journal of Economics*, 123, 905–950.
- WARNEKEN, F. AND M. TOMASELLO (2006): “Altruistic helping in human infants and young chimpanzees,” *Science*, 311, 1248–1249.
- YOUNG, H. P. (2008): “Social norms,” in *The New Palgrave Dictionary of Economics*, ed. by S. N. Durlauf and E. B. Lawrence, Palgrave Macmillan.
- ZUMBUEHL, M., T. J. DOHMEN, AND G. A. PFANN (2013): “Parental investment and the intergenerational transmission of economic preferences and attitudes,” *SOEP Paper n. 570*.

A Online Appendix

A.1 Additional Information

Actions in the Different Treatments Table A1 summarizes the sequence of actions in the different treatments.

Table A1: Treatments

Step	Violation	Violation + Help	Help
1	Actor approaches targeted parent from the front	Actor approaches targeted parent from the front	Actor approaches targeted parent from the front
2	Actor pauses while going through the bag	Actor pauses while going through the bag	Actor pauses while going through the bag
3	Actor litters	Actor litters	-
4	-	Actor takes out bag contents	Actor takes out bag contents
5	Actor continues moving	Actor continues moving	Actor continues moving
6	Actor pauses a second time	Actor drops bag contents on the floor	Actor drops bag contents on the floor
7	Actor leaves the scene	Actor picks up stuff and leaves the scene	Actor picks up stuff and leaves the scene
N_{Child}	100	100	100
N_{Alone}	100	100	101

Power Analysis and Number of Observations We determined an objective of 100 observations per cell and 600 observations in total. Due to the limited literature on this topic, we had no well-established priors with respect to the treatment effect. In comparing Violation scenes across conditions, we drew from the punishment rate in the BasePun treatment of Balafoutas et al. (2014) (17%). With a chi-squared test, we can detect an effect ratio of 2 (i.e., the ratio between the treated-group proportion and the control-group proportion) given the specified sample size, 80% power and a 5% significance level.¹³ A similar analysis for the Help scenes yielded a minimum detectable effect ratio of 1.50 ($w=0.4$), assuming that the baseline helping rate is 39.7% like in the BaseHelp treatment of Balafoutas et al. (2014). Moreover, the helping rate in the HelpViolator treatment of Balafoutas et al. (2014) was 18.6%, which implied a power of 91.31% with our sample size when comparing it to their BaseHelp rate to identify indirect punishment.

Locations We randomly selected 30 out of the 80 public elementary schools in the city of Lyon, the third one in France in terms of population size. We excluded private schools to avoid unobservable selection effects that would possibly interact with our research question. For example, parents using private schools may have different income levels than parents using public schools. By contrast, the assignment of a child to a given public school is determined strictly by the parents' address. Therefore, the average income in the school district gives us indirect information about the wealth of

¹³This corresponds to a medium-large effect size ($w=0.45$).

parents involved in our experiment. In order to make sure that we obtained a representative sample of elementary schools in Lyon, we collected basic information on all public elementary schools in the city, including name and address. We matched each school with the median disposable income and poverty rate of the IRIS area it is located in, and classified each IRIS area as above or below the city-wide median. IRIS are infra-municipal areas comprising between 1800 and 5000 residents. IRIS is an acronym of ‘aggregated units for statistical information’. France is composed of around 16,100 IRIS. We extracted data from the 2014 edition of the INSEE survey “Revenus, pauvreté et niveau de vie”, available at <https://www.insee.fr/fr/statistiques/3288151>. The poverty rate is measured as the share of households with a disposable income below 60 percent of the median income in the city.

We randomly picked schools such that for each socio-economic measure roughly half of the selected schools are classified as above the median, and half are classified as below the median. For the selected schools, we inspected the surroundings in-person and using Google Maps. If the surroundings were not suitable for the scenes (open terrain, steep hills, construction work, etc.), we dropped the school and randomly selected another one to replace it. The resulting diversity in neighborhoods allows us to assess whether the intensity of socialization efforts are affected by the socio-economic environment, as suggested by previous research (Benenson et al., 2007; Bauer et al., 2014; Angerer et al., 2015; Deckers et al., 2017; Kosse et al., 2020).

The vast majority of schools included in the sample host both a kindergarten and a primary school. As a result, the children involved in our experiment are between 3 and 12 years old.¹⁴ We never ran the experiment during more than one day in each school to avoid being identified. We visited each school during one morning and one afternoon on the same day, to avoid targeting the same parents more than once. Thus, the experiment was run on 30 days in total and on average, we collected 20 observations per school. We only ran the experiment in the morning on Wednesdays, due to the fact that schools finish around noon on that day of the week. In addition, we conducted three sessions either only in the morning or in the afternoon. This happened when not everyone from the research team was available the whole day.

A.2 Instructions of the Experiment

[Translated from French]

A.2.1 Instructions for the RAs and the Actor

You are helping us to collect data for an ongoing research project. None of you are aware of the research goal and topic. We work in research teams of four: two RAs, one actor and one supervisor (who is one of the researchers).

We are asking you to stage a number of scenes in the streets around different elementary schools in Lyon. We are interested in the response of the witness of the scene. This witness is an unaware passerby who is targeted by you. We want you to target two types of witnesses: a parent accompanied by one or multiple children, and single

¹⁴We control for the child’s age in the analysis contained in Appendix A.4.6.

adults (who can also be presumed to be parents). You play these scenes in the morning and the afternoon. In each of these time slots you should target both types of adults. This basically means that you target parents with a child going to school, parents leaving school without a child, parents approaching school without their child, and parents leaving school with their child. You should aim for roughly equal numbers in each of these categories.

There are three different scenes to be played, which are further explained below. Below, you find separate instructions for the actor and the RA. Make sure to read each other's instructions, such that you both know what we expect from each of you.

A.2.2 Instructions for the Actor

Materials

- Plain clothes
- Cotton shoulder/shopping bag
- Plastic bag with a banana peel inside
- 7 folders and binders
- Colored pens and markers
- Two tablets for RAs

Before the scene is played, together with the RAs and supervisor, you determine the location of the scene. Make sure not to be too close to the school entrance, in a street that is not too busy. Your first task is to identify your target. You should target either a single adult or a single adult accompanied by children. Do not target adults with a stroller, a bike, a dog, parents of a disabled child, or parents who are holding their children's hands with both hands. Also make sure to avoid parents visibly in a rush or talking on the phone.

For each targeted adult, you play one of three scenes. Before the scenes, make sure to have one handle of the bag on your shoulder and the other loosely hanging down; this makes it easier to reach into your bag. Below is a detailed script of the scenes.

Scene 1: Violation + Help

1. Actor and supervisor identify targeted parent fulfilling the criteria.
2. Actor approaches target from the front.
3. As the actor is roughly 10 meters away, (s)he pauses and pretends to be searching for something in his/her bag.
4. As the target is 5 meters away: the actor throws away the plastic bag with the banana peel inside. The actor makes sure that no one is approaching from behind.
5. Actor takes out all file folders from the bag and starts moving again. As (s)he continues to walk, the actor accidentally drops the entire content.

6. Actor stops walking, reacts visibly upset, stares at dropped items in dismay. This provides the parent with an opportunity to help.
7. Actor starts picking up items as targeted parent passes him/her.
8. Scene ends; RA and Actor record information and clean up scene.

Scene 2: Violation

1. Actor and supervisor identify targeted parent fulfilling the criteria.
2. Actor approaches target from the front.
3. As the actor is roughly 10 meters away, (s)he pauses and pretends to be searching for something in his/her bag.
4. As the target is 5 meters away: the actor throws away the plastic bag with the banana peel inside. The actor makes sure that no one is approaching from behind.
5. Actor starts moving again, but, before the target has reached him/her, then pauses again, going through the bag again as the parent passes.
6. Scene ends; RA and Actor record information and clean up scene.

Scene 3: Help

1. Actor and supervisor identify targeted parent fulfilling the criteria.
2. Actor approaches target from the front.
3. As the actor is roughly 10 meters away, (s)he pauses and pretends to be searching for something in his/her bag.
4. Actor takes out all file folders from the bag and starts moving again. As (s)he continues to walk, the actor accidentally drops the entire content.
5. Actor stops walking, reacts visibly upset, stares at dropped items in dismay. This provides the parent with an opportunity to help.
6. Actor starts picking up items as targeted parent passes him/her.
7. Scene ends; RA and Actor record information and clean up scene.

After the end of each of the scenes, you leave the location in a direction different from that of the targeted adult. In case that the adult confronts you about throwing away the plastic bag and/or demands you clean it up (in scenes 1 and 2), you quietly comply and pick up the plastic bag. If the parent does not respond, the RA makes sure to clean up after the parent has left the scene. After the scene, you meet with the RA and report the following pieces of information:

- Type of scene played
- Reaction of the parent (multiple could apply):

- Punishment: the parent explicitly addresses you regarding the littering and/or demands you to clean it up.
 - Help: the parent picks up at least one of the dropped item from the floor.
 - Address child: the parent talks to the child about the violation in a way that is audible for you.
- “Other” circumstances

A.2.3 Instructions for Research Assistant 1

Each team includes two RAs, each with different tasks. RA1 observes the scene played closely and notes down the following characteristics:

- ID: School code + number of observation, e.g., GT11 for the 11th observation at Germaine Tillion.
- Setting: what are the circumstances in which the scene is played?
 - Time of day
 - Witnesses: are there any other people around that could possibly intervene in the scene?
 - Weather: sunny, cloudy, or rainy; cold, mild, hot; windy?
 - Direction: from or to school
 - Cleanliness of environment: scale of 1 (dirty) to 5 (clean)
- Treatment: confirm this with the actor after the scene.
 - Condition: child(ren) or alone
 - Type of scene: Violation+Help, Violation, Help
 - In case of littering: did the target see the plastic bag being thrown away?
- Reaction of the parent: confirm this with the actor after the scene.
 - Punishment: does the parent confront the actor by directly addressing him/her about the violation?
 - Help: does the parent help by picking up at least one item?
 - Address child: does the parent talk to the child about the violation?

It is important that RA1 does not stay too close to the parent, because this may contaminate the outcome of the scene. After the end of the scene, RA1 verifies the scene played with the actor and checks whether the parent said something to the actor.

A.2.4 Instructions for Research Assistant 2

The task of RA2 is to approach the target after the scene for a seemingly unrelated survey. You tell the target the following:

[Translated from French]

“Good day sir/madam, I am a Master student in Psychology at the University of Lyon 2 and, as part of my courses, I am conducting a survey on the quality of the environment around schools. The survey comprises 5 questions and takes 2 minutes. Could I take some of your time to respond to my questions?”

1. We are close to the elementary school [name of school]. How would you evaluate the quality of the air around this school on a scale of 1 (for a very poor quality) to 7 (for an excellent quality)?
2. Do you think that the circulation of cars should be forbidden in the streets in front of schools to limit the exposure to pollution for children?
3. Today, are you accompanying or have you accompanied your child / a child that you guard / a child of one of your relatives / or no child to this school?
4. If so, what is the age and gender of this child / these children?
5. Finally, do you take the car to arrive at school?

You should also note down a number of characteristics regarding the parent’s appearance. The main purpose for this is to ensure that no parent is targeted twice.

- ID: School code + number of observation, e.g., GT11 for the 11th observation at Germaine Tillion.
- Gender: male or female
- Estimated age
- Ethnicity: caucasian, Arab, African, Asian, other (Indian, South-American)
- Religious signs
- Estimated height
- Build: lean, medium, overweight, obese.
- Hair colour: blond, light brown, dark brown, black, red, gray, other
- Hair style: bold, short, medium, long, curly, straight, ponytail, afro.
- Facial hair: none, moustache, short beard, long beard
- Colour of outer garment (coat, vest, etc.)
- Other: hat, glasses, tattoos, piercings, birth marks, scars, etc.

A.3 Additional Materials

Figure A.1: Materials Used in the Experiment and Scenes



A.4 Supplementary Analyses

A.4.1 Response Rates and Summary Statistics of Survey Respondents

Table A2 reports the response rates to the survey in the morning and in the afternoon, by condition.

Table A2: Response Rates to the Survey, by Condition

	MORNING				AFTERNOON			
	ALONE	CHILD	ALL	χ^2 test	ALONE	CHILD	ALL	χ^2 test
Help	0.48 (0.07)	0.60 (0.07)	0.54 (0.05)	0.221	0.46 (0.08)	0.62 (0.08)	0.54 (0.06)	0.162
Violation	0.32 (0.07)	0.45 (0.07)	0.39 (0.05)	0.212	0.40 (0.08)	0.51 (0.09)	0.46 (0.06)	0.337
Violation+Help	0.44 (0.08)	0.39 (0.07)	0.41 (0.05)	0.624	0.40 (0.07)	0.55 (0.09)	0.46 (0.06)	0.200
All scenes	0.41 (0.04)	0.48 (0.04)	0.45 (0.03)	0.303	0.42 (0.04)	0.56 (0.05)	0.48 (0.03)	0.032

Note: The “ χ^2 test” columns display the p-values from Pearson’s chi-squared tests comparing response rates between the Alone and Child conditions.

Table A2 shows that regardless of the time of the day, parents were consistently more likely to respond to the survey when the child was present than when they were alone in all treatments, except in the Violation+Help treatment in the morning.¹⁵ They also tend to respond more to the survey after observing the actor needing help (in the Help treatment) than after observing a norm violation (χ^2 tests: Help vs. Violation, $p = 0.026$; Help vs. Violation+Help, $p = 0.046$), which suggests that there might be some spillover effects of the scene on the survey response rate. These differences in response rates across treatments are not consistent with an interpretation of our main results being driven by time pressure. Indeed, time pressure should be orthogonal to the treatment. We also note that if parents were in a hurry in any condition, the response rates to the survey would have probably been much lower.¹⁶

Additionally, we explored the correlation between responding to the survey and helping in the Help treatment, depending on whether the child was present or not, by means of linear regressions. The idea was to test whether the two helping behaviors were complements or substitutes, and whether this was depending on the presence of the child. However, we failed finding significant correlations (Child condition: coefficient = -0.13, $p=0.242$; Alone condition: coefficient = 0.07, $p=0.585$).

Table A3 presents summary statistics on survey participants. There are no clear significant differences between the Child and Alone conditions regarding the number of

¹⁵The difference in response rates between Child and Alone is statistically significant at the aggregate level (χ^2 test, $p = 0.035$) and in the afternoon (χ^2 test, $p = 0.032$) if we pool all scenes together.

¹⁶As a further robustness check of the importance of time pressure, we ran the same regressions of Table 2 but only on the sample of subjects who could be reached for the survey. Among the explanatory variables, we included a dummy equal to 1 if a subject responded to the survey, and 0 otherwise. The idea was to use the response to the survey as a proxy for *not* being in a rush. Our main results did not change qualitatively. In addition, the coefficient of the survey dummy turned out not significant. The results are available upon request.

children and their age.

Table A3: Summary Statistics of Survey Respondents

	(1) ALL	(2) ALONE	(3) CHILD	(4) DIFF.
Own Child	0.88 (0.33)	0.91 (0.29)	0.85 (0.36)	0.06 (0.05)
Son	0.61 (0.49)	0.62 (0.49)	0.61 (0.49)	0.01 (0.07)
Daughter	0.64 (0.48)	0.63 (0.49)	0.64 (0.48)	-0.01 (0.07)
No. of Children	1.46 (0.64)	1.40 (0.59)	1.52 (0.67)	-0.12 (0.09)
Child Age = 3	0.06 (0.25)	0.06 (0.24)	0.06 (0.25)	-0.00 (0.03)
Child Age = 4	0.16 (0.37)	0.16 (0.37)	0.17 (0.38)	-0.01 (0.05)
Child Age = 5	0.26 (0.44)	0.25 (0.44)	0.27 (0.45)	-0.02 (0.06)
Child Age = 6	0.18 (0.39)	0.17 (0.38)	0.19 (0.40)	-0.03 (0.05)
Child Age = 7	0.23 (0.42)	0.19 (0.39)	0.26 (0.44)	-0.07 (0.06)
Child Age = 8	0.21 (0.41)	0.23 (0.42)	0.20 (0.40)	0.03 (0.06)
Child Age = 9	0.19 (0.39)	0.15 (0.36)	0.22 (0.41)	-0.07 (0.05)
Child Age = 10	0.08 (0.27)	0.11 (0.31)	0.06 (0.23)	0.05 (0.04)
Child Age = 11	0.03 (0.18)	0.03 (0.18)	0.03 (0.18)	-0.00 (0.02)

Notes: For gender and age, totals are not equal to 1 because some parents reported having more than one child at the school. * $p < 0.10$.

A.4.2 Regression models with Child \times VH interaction

Table A4 reproduces the regressions reported in the main paper but with the inclusion of an interaction term between the Child and Witness variables. The main effects are equivalent to those reported in the main paper. The coefficient of the Witness dummy (that captures the pure presence of another adult observing) is negative and significant in model (1), and negative but insignificant in model (2): being observed by another adult reduces the parents' likelihood of punishing the violator and does not affect the willingness to help. The interaction term between the Child and Witness variables is significant in neither regression: parents do not enforce more the norm when the child is present because they are observed by another adult.

A.4.3 Effect of Parent's Gender on Punishment in the Violation Treatment

To explore the role of the parent's gender on punishment behavior, Table A5 shows an additional analysis of the punishment rate in the Violation treatment only. When the Child \times Male Target interaction is added, Child alone is no longer significant. At the same time, Child and Child \times Male Target are jointly significant ($p = 0.046$). This means that only fathers punish significantly more in the presence of the Child. We take this as indirect evidence that the presence of the child raises the fear of retaliation, but that fathers experience this fear less. This is strengthened by the fact that, in the

Table A4: Determinants of Punishment Rate (Left) and Helping Rate (Right)

	(1)	(2)
Child	0.11** (0.04)	0.22*** (0.05)
VH	-0.04 (0.03)	-0.10** (0.04)
Child × VH	-0.13*** (0.04)	-0.07 (0.08)
Male Target	0.06** (0.03)	0.03 (0.04)
Male Actor	-0.06 (0.04)	-0.22*** (0.05)
Morning	-0.04 (0.03)	0.01 (0.04)
Witness	-0.09*** (0.03)	-0.15 (0.10)
Child × Witness	0.09 (0.07)	0.02 (0.10)
Rich area	0.02 (0.04)	0.01 (0.05)
Rain	0.01 (0.03)	-0.08 (0.07)
Hot	-0.03 (0.04)	0.06 (0.05)
Constant	0.14*** (0.05)	0.37*** (0.05)
Observations	399	400
Clusters	30	30
R^2	0.07	0.13
F	6.45	6.00
df	29	29

Notes: The table contains results from pooled Ordinary Least Squared regressions. The dependent variable is a binary variable for punishment (columns 1) or helping (columns 2). Standard errors in parentheses are clustered at the school level (30 clusters). One observation is dropped due to missing data on the target’s gender. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

absence of the Child, fathers only punish insignificantly more, indicating that the fear of retaliation is much more similar between mothers and fathers when the Child is not around.

A.4.4 Alternative Estimation Models

Instead of linear probability models, we also estimated logit models. While such models are more suited to analyze binary choice data like ours, they are less suited to study interaction effects. To this end, we took the following approach. We estimated the logit model including the controls and the interaction term. We then estimated marginal effects at Child = 1 and Child = 0 and used a contrast test to determine whether the marginal effects of VH are significantly different between the conditions. These results are presented in Table A6. As can be seen, the marginal effect of VH differs regarding Punishment, but not for Helping.

Table A5: Punishment in the Violation Treatment

	(1)	(2)
Child	0.11** (0.04)	0.09 (0.06)
Male Target	0.10** (0.04)	0.06 (0.06)
Child \times Male Target		0.07 (0.15)
Male Actor	-0.09* (0.05)	-0.09* (0.05)
Morning	-0.02 (0.05)	-0.03 (0.05)
Witness	-0.10 (0.08)	-0.10 (0.07)
Rich area	0.06 (0.05)	0.06 (0.06)
Rain	-0.00 (0.06)	-0.00 (0.06)
Hot	-0.09 (0.07)	-0.09 (0.07)
Constant	0.14* (0.07)	0.15* (0.08)
Observations	199	199
Clusters	29	29
R^2	0.07	0.07
F	4.47	5.48
df	28	28

Notes: Standard errors are clustered on the School level. Only the Violation treatment is included. One observation is dropped due to missing data on the target’s gender. Male Target and Child \times Male Target are jointly significant ($F(2, 28) = 3.43, p = 0.046$). ** $p < 0.05$.

A.4.5 Sample Restrictions

Below, we show our results to be robust to a number of sample restrictions and alternative definitions. We summarize the results of this endeavor in Table A7, building upon the pooled linear probability models (2) and (4) of Table 2. Column (1) and (4) of Table A7 show that our results on punishment and helping, respectively, are unaffected by excluding targets who were identified as guardians and not as parents through the survey. In columns (2) and (5), we only include parents accompanying at most one child for whom helping could have been easier compared to parents accompanying several children. The coefficient on $VH \times Child$ for Helping remains insignificant. Then, in columns (3) and (6), we discard observations for which a witness was recorded by the RA. Again, this does not change the previous results. Finally, in column (7), we re-coded the Helping dummy so that a child encouraged by the parent to provide help is now coded as *no* Helping. Since this only happened in 8 instances (5 in Help and 3 in Violation + Help), it does not change the estimates substantially.

Table A6: Marginal Effects of Logit Estimations

	PUNISHMENT		HELPING	
	ALONE (1)	CHILD (2)	ALONE (3)	CHILD (4)
Child	0.12*** (0.04)	0.12*** (0.04)	0.22*** (0.05)	0.22*** (0.05)
VH	-0.04 (0.03)	-0.16*** (0.04)	-0.10** (0.04)	-0.17** (0.07)
Male Target	0.07* (0.04)	0.12*** (0.05)	0.03 (0.04)	0.04 (0.05)
Male Actor	-0.07 (0.05)	-0.13* (0.07)	-0.21*** (0.04)	-0.28*** (0.06)
Morning	-0.04 (0.03)	-0.08 (0.05)	0.00 (0.04)	0.00 (0.05)
Witness	-0.04 (0.04)	-0.08 (0.06)	-0.11* (0.06)	-0.16* (0.09)
Rich area	0.03 (0.04)	0.05 (0.07)	0.01 (0.05)	0.02 (0.06)
Rain	0.00 (0.03)	0.00 (0.06)	-0.10 (0.07)	-0.14 (0.10)
Hot	-0.04 (0.04)	-0.07 (0.07)	0.05 (0.05)	0.06 (0.06)
VH _C vs. VH _A	9.94***		0.79	
Observations	399		400	
Clusters	30		30	
(Pseudo) R^2	0.097		0.113	
Wald χ^2	55.93		52.45	

Notes: The table contains four sets of marginal effects resulting from two logit estimations: one with Punishment as the dependent variable (columns 1 and 2) and one with Helping as the dependent variable (columns 3 and 4). For each estimation, marginal effects are estimated for Child = 0 and Child = 1. Delta-method standard errors are reported in parentheses. The row “VH_C vs. VH_A” displays the χ^2 -test statistic of a contrast test against the null that the coefficients on VH are the same in the two conditions. In the regression with Punishment as the dependent variable, one observation is dropped due to missing data on the target’s gender. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

A.4.6 Analysis of Heterogeneous Effects

In this part of the Appendix, we report additional regressions in which we allow the presence of the child to have heterogeneous effects with regards to different characteristics of the child, the parent, or the neighborhood. The results of this endeavor are depicted in Table A8 for Punishment and Table A9 for Helping. We discuss the tables jointly, as each column in the two tables corresponds to the *same* exercise. Column (1) contains the baseline estimates, as reported in models (2) and (4) in Table 2 in the main text with controls included.

Number of Children First, we look at the importance of the number of children accompanying the parent. In column (2) in both tables, we discriminate between parents accompanying one child, two children, and three or more children. It should be noted that only 17 out of 300 parents in the Child condition accompany three or more children

Table A7: Robustness Checks

	PUNISHMENT			HELPING			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Child	0.13*** (0.05)	0.13** (0.06)	0.11** (0.04)	0.19*** (0.06)	0.18** (0.07)	0.19*** (0.06)	0.17*** (0.05)
VH	-0.05* (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.11** (0.04)	-0.10** (0.04)	-0.13** (0.06)	-0.10** (0.04)
Child × VH	-0.13** (0.05)	-0.13** (0.06)	-0.14*** (0.05)	-0.04 (0.08)	-0.11 (0.10)	-0.01 (0.08)	-0.05 (0.07)
Constant	0.14*** (0.05)	0.13** (0.05)	0.16*** (0.05)	0.39*** (0.05)	0.37*** (0.05)	0.37*** (0.06)	0.36*** (0.05)
Controls	✓	✓	✓	✓	✓	✓	✓
Observations	385	319	347	383	318	334	400
Clusters	30	30	30	30	30	30	30
R^2	0.08	0.07	0.07	0.13	0.11	0.13	0.10
F	4.91	2.78	3.92	6.21	7.18	7.41	5.61
df	29	29	29	29	29	29	29

Notes: Dependent variable is a dummy for Punishment (columns (1) to (3)) or Helping (columns (4) to (7)). Standard errors in parentheses are clustered at the school level. Columns (1) and (4) present regression results when identified guardians are excluded from the analysis. In columns (2) and (5), we exclude all parents accompanying 2 or more children. Columns (3) and (6) include only observations for which no witness was recorded. Finally, in column (7) we code the child(ren) helping as *not* helping, rather than helping. These regressions control for the same variables as models (2) and (5) in Table 2, but the coefficients are omitted here for the sake of concision. One observation is dropped due to missing data on the target’s gender. * $p < 0.10$ ** $p < 0.05$ *** $p < 0.01$.

(2 in Violation, 9 in Help, 6 in Violation + Help treatments). Parents who accompany three or more children punish significantly less than parents alone (at the 1% level). This may suggest that parents accompanying three or more children are too occupied paying attention to the children or consider that the children are less attentive to what they would be doing to engage in punishment. A slightly different picture arises when looking at Helping. The helping rate is significantly higher for all numbers of children. However, we see that only parents accompanying one child withhold helping substantially (by 12 pp), even though the coefficient enters insignificantly.

Child’s Gender Next, we investigate whether the child’s gender matters in the parent’s reaction (see column (3) in both tables). To allow for a clean comparison, only single-child observations are classified according to gender. In total, 22 single-child observations have missing gender of the child and are omitted from the analysis. The results show the presence of one girl raises the punishment rate by 18 percentage points, while the presence of one boy raises it by 12 points. This coefficient is borderline significant for the presence of one girl, only. Additionally, the presence of one girl raises the helping rate by 17 percentage points, while the presence of one boy raises it by 15 points. The coefficient is borderline significant both for girls and boys. For both genders, the additional drop in helping in the VH treatment is insignificant and of similar magnitude. Taken together, parents’ educative motive is not really stronger with daughters than with sons.

Table A8: Secondary Analyses of Punishment Behavior

	BASE (1)	CHLD (2)	CHLD (3)	CHLD (4)	TARGET (5)	TARGET (6)	INCOME (7)	INCOME (8)
Child	0.12** (0.04)				0.07 (0.05)	0.12** (0.04)	0.15*** (0.05)	0.09*** (0.02)
1 Child		0.13* (0.06)						
1 Boy			0.12 (0.11)					
1 Girl			0.18* (0.09)					
1 Child Age ≤ 5				0.07 (0.13)				
5 < 1 Child Age ≤ 8				0.19* (0.09)				
1 Child Age > 8				0.01 (0.10)				
2 Children		0.11* (0.06)	0.11* (0.06)	0.12* (0.06)				
3+ children		-0.14*** (0.05)	-0.15*** (0.05)	-0.15*** (0.05)				
VH	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.05* (0.03)
VH × Child	-0.12*** (0.04)				-0.12*** (0.04)	-0.12*** (0.04)	-0.12** (0.05)	-0.12** (0.05)
VH × 1 Child		-0.13** (0.06)						
VH × 1 Boy			-0.09 (0.12)					
VH × 1 Girl			-0.23** (0.09)					
VH × Child Age ≤ 5				-0.11 (0.14)				
VH × 5 < Child Age ≤ 8				-0.16* (0.10)				
Child Age > 9				-0.06 (0.11)				
VH × 2 Children		-0.11 (0.08)	-0.12 (0.08)	-0.11 (0.08)				
VH × 3+ Children		0.09 (0.06)	0.10 (0.06)	0.10 (0.06)				
Male Target	0.06** (0.03)	0.07** (0.03)	0.06** (0.03)	0.06** (0.03)	0.05 (0.04)	0.08* (0.05)	0.07** (0.03)	0.07** (0.03)
Male Target × Child					0.02 (0.08)			
Male Actor	-0.06 (0.04)	-0.06 (0.04)	-0.06 (0.04)	-0.06 (0.04)	-0.10** (0.05)	-0.05 (0.04)	-0.06 (0.04)	-0.03 (0.03)
Male Actor × Child					0.07 (0.06)			
Male Actor × Male target						-0.03 (0.05)		
Rich IRIS	0.02 (0.04)	0.03 (0.04)	0.04 (0.04)	0.04 (0.04)	0.02 (0.04)	0.03 (0.04)	0.06 (0.05)	
Rich IRIS × Child							-0.07 (0.06)	
Medium Low Income								0.14*** (0.04)
Medium High Income								0.08** (0.04)
High Income								0.17** (0.06)
Medium Low Income × Child								0.09 (0.07)
Medium High Income × Child								0.04 (0.03)
High Income × Child								-0.03 (0.07)
Morning	-0.04 (0.03)	-0.05 (0.03)	-0.04 (0.04)	-0.04 (0.04)	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)	-0.04 (0.03)
Witness	-0.04 (0.04)	-0.04 (0.04)	-0.05 (0.05)	-0.08 (0.05)	-0.04 (0.04)	-0.04 (0.04)	-0.03 (0.04)	-0.02 (0.04)
Rain	0.01 (0.04)	0.02 (0.03)	0.03 (0.03)	0.03 (0.04)	0.01 (0.04)	0.01 (0.03)	0.01 (0.04)	-0.01 (0.03)
Hot	-0.04 (0.05)	-0.04 (0.05)	-0.05 (0.05)	-0.04 (0.05)	-0.04 (0.05)	-0.04 (0.05)	-0.04 (0.05)	-0.05 (0.03)
Constant	0.14*** (0.05)	0.14*** (0.05)	0.13*** (0.05)	0.13*** (0.05)	0.16*** (0.05)	0.13*** (0.04)	0.12** (0.04)	0.03 (0.03)
Observations	399	399	384	384	399	399	399	399
Clusters	30	30	30	30	30	30	30	30
R ²	0.07	0.07	0.09	0.09	0.07	0.07	0.07	0.10
F	5.17	9.87	12.34	9.55	5.41	4.66	5.12	13.36
df	29	29	29	29	29	29	29	29

Notes: Standard errors in parentheses are clustered at the School level. * $p < 0.1$
** $p < 0.05$ *** $p < 0.01$.

Table A9: Secondary Analyses of Helping Behavior

	BASE (1)	CHILD (2)	CHILD (3)	CHILD (4)	TARGET (5)	TARGET (6)	INCOME (7)	INCOME (8)
Child	0.22*** (0.06)				0.28*** (0.08)	0.22*** (0.06)	0.16* (0.09)	0.22*** (0.09)
1 Child		0.19** (0.07)						
1 Boy			0.15* (0.08)					
1 Girl			0.17* (0.10)					
1 Child Age ≤ 5				0.26** (0.12)				
5 < 1 Child Age ≤ 8				0.14 (0.12)				
1 Child Age > 8				0.05 (0.14)				
2 Children		0.27*** (0.09)	0.28*** (0.09)	0.28*** (0.09)				
3+ children		0.26** (0.10)	0.27** (0.10)	0.26** (0.10)				
VH	-0.10** (0.04)	-0.10** (0.04)	-0.10** (0.04)	-0.10** (0.04)	-0.11** (0.04)	-0.10** (0.04)	-0.10** (0.04)	-0.11** (0.04)
VH × Child	-0.07 (0.08)				-0.06 (0.08)	-0.07 (0.08)	-0.08 (0.07)	-0.07 (0.08)
VH × 1 Child		-0.12 (0.10)						
VH × 1 Boy			-0.15 (0.12)					
VH × 1 Girl			-0.08 (0.11)					
VH × 1 Child Age ≤ 5				-0.26 (0.17)				
VH × 5 < 1 Child Age ≤ 8				-0.07 (0.15)				
1 Child Age > 9				0.05 (0.16)				
VH × 2 Children		-0.02 (0.14)	-0.03 (0.14)	-0.03 (0.14)				
VH × 3+ Children		0.03 (0.24)	0.03 (0.24)	0.03 (0.24)				
Male Target	0.03 (0.04)	0.02 (0.04)	0.02 (0.04)	0.02 (0.04)	0.05 (0.06)	0.05 (0.06)	0.03 (0.04)	0.03 (0.04)
Male Target × Child					-0.06 (0.11)			
Male Actor	-0.22*** (0.05)	-0.23*** (0.05)	-0.22*** (0.05)	-0.22*** (0.05)	-0.19*** (0.06)	-0.21*** (0.05)	-0.22*** (0.05)	-0.20*** (0.05)
Male Actor × Child					-0.07 (0.09)			
Male Actor × Male target						-0.05 (0.07)		
Rich IRIS	0.01 (0.05)	0.02 (0.05)	0.02 (0.05)	0.02 (0.05)	0.01 (0.05)	0.01 (0.05)	-0.06 (0.06)	
Rich IRIS × Child							0.14 (0.09)	
Medium Low Income								0.08 (0.10)
Medium High Income								-0.00 (0.10)
High Income								0.08 (0.07)
Medium Low Income × Child								-0.02 (0.15)
Medium High Income × Child								0.02 (0.13)
High Income × Child								0.02 (0.09)
Morning	0.01 (0.04)	0.02 (0.04)	0.02 (0.04)	0.02 (0.04)	0.01 (0.04)	0.01 (0.04)	-0.00 (0.04)	-0.00 (0.04)
Witness	-0.13 (0.08)	-0.13 (0.08)	-0.16* (0.08)	-0.16** (0.08)	-0.13 (0.08)	-0.14* (0.08)	-0.14* (0.08)	-0.13 (0.08)
Rain	-0.08 (0.07)	-0.08 (0.07)	-0.08 (0.07)	-0.08 (0.08)	-0.08 (0.08)	-0.09 (0.07)	-0.09 (0.08)	-0.08 (0.08)
Hot	0.06 (0.05)	0.06 (0.05)	0.06 (0.06)	0.06 (0.06)	0.06 (0.05)	0.06 (0.05)	0.06 (0.05)	0.06 (0.05)
Constant	0.36*** (0.05)	0.36*** (0.05)	0.36*** (0.05)	0.36*** (0.05)	0.34*** (0.06)	0.36*** (0.05)	0.40*** (0.06)	0.32*** (0.08)
Observations	400	400	388	390	400	400	400	400
Clusters	30	30	30	30	30	30	30	30
R ²	0.13	0.14	0.14	0.15	0.13	0.13	0.14	0.14
F	6.45	4.81	4.65	4.33	5.76	5.76	15.41	5.98
df	29	29	29	29	29	29	29	29

Notes: Standard errors in parentheses are clustered at the School level. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$.

Child's Age Then, we explore the effect of the age of the child (see column (4)). Again, we do this by focusing on single-child observations for the cleanest comparison. We created dummies for one child aged 5 or younger, one child aged between and including 6 and 8, and one child aged 9 or older. The values of this variable are based either on the parent's response in the survey or on the guess of the research assistants (21 observations are missing).

The results show that the increase in Punishment in the presence of the child seems to be driven by the middle age category, as the increase in punishment rates is (weakly) significant only for parents accompanying children aged between 6 and 8 at a magnitude of 19 percentage points. For the youngest category, the coefficient is closer in magnitude to the baseline estimate, but insignificant, while the coefficient is close to zero for the oldest category. A somewhat different picture arises for helping, as parents accompanying the youngest class of children respond strongest to the presence of the child by increasing the helping rate by 26 percentage points (significant at 5%-level). Parents of older children increase their helping rate by less as compared to parents alone, and these increases are insignificant. Moreover, the additional drop in helping rate is large (though not significant) for the youngest class of children, only. It thus seems that results regarding Helping are mainly driven by the youngest age category. These contrasting results paint an interesting picture. Parents seem to focus on modeling good behavior, that is, helping a stranger in need, to younger children, and become more sophisticated in their teaching of norms by punishing norm violations of others when children become older.

Targeted Parent and Actor Gender Fathers and mothers may react differently and they may also react to the gender of the actor in the presence of the child. In column (5) and (6) we look at the effects of the gender of the targeted parent. Most importantly, we see no interaction effects between the gender of the parent or actor and the presence of the child, as indicated by the insignificant coefficients on the interaction terms. Interestingly, the coefficient on Child now enters insignificantly in the Punishment regression as it now refers to the presence of the Child with a Female target and the Female actor. Furthermore, the interaction term Male Actor \times Male Target measures the effect of two males interacting. This does not seem to affect outcomes significantly. More generally, most regressions show that punishment does not differ according to the actor's gender (except in column (5)), while the male actor receives significantly less help than the actress regardless of the specification.

Income Effects Finally, we would like to know whether income influences parents' punishment and helping. To this end, in column (7) we interact the Rich IRIS dummy with the Child dummy. We find no effects of this interaction term for both outcomes. To dive deeper into this, in column (8) we classify the IRIS area in which the school is located as Low, Medium Low, Medium High, or High based on the median disposable income. The results show that parents in the highest three income classes punish significantly more than parents in the lowest income class. However, parents do not increase their punishment by more in these neighborhoods in the presence of the child. Regarding helping rates, we do not find effects of income on the parents' tendency to provide help. Additional regressions in which we include the poverty rate in the IRIS

area and an interaction term with the Child variable lead to similar conclusions. Parents are significantly less likely to punish in poorer areas (at the 5% level) but do not help less, and the effect of the presence of the child does not differ significantly with the poverty rate (regressions available upon request).

A.4.7 Analysis of Timing Effects

Throughout the experiment, parents arrive at different times at school. It may be that parents who arrive early are different from parents arriving later, either because they are less in a rush or because they are intrinsically different. Similarly, parents who leave the school premises late may be different from parents leaving the premises as soon as possible. If this is the case, this may affect their punishment or helping behavior. In order to test whether this is the case, we take three approaches and report the results in Table A10. The first approach uses the timing of the scene in minutes relative to the beginning or end of the school day (8:30 AM, 4:45 PM, or 5:30 PM). The second approach rounds the previous time variable to the nearest ten, in order to discretize the support. The third approach uses the observation number within a session (morning or afternoon) and condition (*i.e.*, Child or Alone).

Table A10 shows no significant effects of the timing of the scene regardless of the approach retained. This suggests that we succeeded in avoiding parents who are in a rush, either because they are late picking up or dropping off their kid, or because they need to be somewhere after picking up or dropping off their kid.

Nevertheless, being in a hurry may depend on the time it took to drop off the child at school and to pick the child from school, especially if the process took longer than expected. We cannot measure this directly but we can test indirectly whether parents might be more in a hurry after dropping their child (that is, in the Alone condition in the morning) or after picking their child (that is, in the Child condition in the afternoon). We estimated the same models as those reported in Table 2 in the main text after adding an interaction term ‘Morning x Child’. This term is significant neither in the Punishment regression (coefficient = -0.02 , standard error = 0.09), nor in the Helping regression (coefficient = -0.09 , standard error = 0.09). Additionally, if individuals were more in a hurry on their way out of school, we should see that they are less likely to respond to the survey when they are alone in the morning and with their child in the afternoon. This is not the pattern that we observe (see the analysis of response rates to the survey in the Appendix section A.4.1). This again suggests that time pressure was not different across conditions.

Table A10: Role of the Timing of Scenes on Punishment and Helping

	PUNISHMENT			HELPING		
	(1)	(2)	(3)	(4)	(5)	(6)
Child	0.114** (0.049)	0.111** (0.048)	0.113** (0.043)	0.216*** (0.054)	0.206*** (0.054)	0.223*** (0.057)
VH	-0.041 (0.028)	-0.040 (0.028)	-0.039 (0.029)	-0.104** (0.044)	-0.106** (0.044)	-0.096** (0.045)
Child × VH	-0.119*** (0.043)	-0.120*** (0.042)	-0.123*** (0.042)	-0.067 (0.078)	-0.066 (0.079)	-0.070 (0.078)
Time × Arriving	-0.001 (0.002)			-0.000 (0.002)		
Time × Leaving	-0.001 (0.002)			0.000 (0.003)		
T10 × Arriving		-0.001 (0.002)			-0.001 (0.002)	
T10 × Leaving		-0.000 (0.001)			0.001 (0.003)	
No. × Arriving			-0.009 (0.006)			-0.008 (0.011)
No. × Leaving			-0.009 (0.008)			0.005 (0.013)
Constant	0.140*** (0.047)	0.138*** (0.046)	0.169*** (0.050)	0.363*** (0.051)	0.362*** (0.050)	0.368*** (0.065)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	399	399	399	400	400	400
Clusters	30	30	30	30	30	30
R^2	0.07	0.07	0.07	0.13	0.13	0.13
F	4.54	4.38	4.59	7.19	6.97	6.16

Notes: TIME is a variable that measures the number of minutes from the beginning or end of school. Negative values indicate a time *before* the school bell rings, while positive values correspond to a time *after* the school bell has rung. TIME is truncated at -40 and 40 . T10 rounds TIME to the nearest ten, in order to discretize the support. Finally, No. denotes the observation number *within* the same condition and time of day. For example, No. = 2 corresponds to the second observation of a given condition at a given time of day. One observation is dropped due to missing data on the target's gender. ** $p < 0.05$ *** $p < 0.01$.

A.5 A Vignette Study for Norm Elicitation

A.5.1 Do social norms differ in the presence of a child?

In Section 3, we showed that parents exhibit a higher tendency to enforce a social norm when in the presence of a child. In this section and the following one, we aim to distinguish between two competing explanations for these results by means of a norm-elicitation survey. Indeed, on the one hand, the presence of the child may motivate the parent to enforce the social norm more than when the parent would be on his or her own. On the other hand, the presence of the child could *change* the social norm in itself and could make littering a more serious violation in the presence of the child. If this is the case, teaching the child may not be the only motive raising parents’ tendency to punish. To study this, we first conducted a vignette study eliciting the social appropriateness of the violation, direct punishment, *not* helping, and *not* helping after a violation (*i.e.*, indirect punishment), respectively, in the presence and in the absence of a child.

This survey was conducted in two waves three months after the field experiment and it was not pre-registered. Accompanied by a supportive document from the regional school authority inspection, we sent a letter and a poster to the principals of all the public elementary schools in Lyon. In this letter, we asked them to send all parents in their school a link with an invitation to participate in an online survey, and to place the poster at the school. We also contacted them through e-mail asking them to forward our attached message to the parents. In total 506 parents responded to our survey.

In the survey, the respondents had to read vignettes presenting all three treatments of our study in one of the two conditions, Alone or Child (see details in subsection A.5.2 below). The order of the Violation and Help vignettes was randomized, while the Violation + Help vignette was shown last. Thus, the treatments were varied within-subjects, while the condition was varied between-subjects. Respondents were asked to rate the social appropriateness of the described behavior on a six-point scale ranging from “very socially inappropriate” to “very socially appropriate” (a neutral option was omitted) and were told that they would have the chance to win a tablet if they chose the option that was chosen by the majority of other respondents. This way, respondents were incentivized to choose the option that they perceived as the social norm (Krupka and Weber, 2013).

Results are presented in Table A11. We coded the ratings as equidistant values on a range from -1 to 1 , with the former indicating “very socially inappropriate” and the latter “very socially appropriate”. From Panel A, it seems that parents shown the Child condition deemed the littering violation slightly more inappropriate, with an average appropriateness rating of -0.83 , as compared to the Alone condition with an average rating of -0.77 . However, on further inspection, we discovered a potential confound in the description of the violation scenario that was not present in the actual treatments. In particular, the banana peel was made very salient and not said to be contained in a plastic bag, as was the case in the experiment. This may have raised the perceived risk of the child slipping on the banana peel and, as a result, increased the severity of the violation. We therefore adapted the description of the vignettes mid-way to exclude

Table A11: Social Appropriateness of Passerby Behavior

	CHILD CONDITION							ALONE CONDITION							Rank-sum test (z)
	Mean	-1	-0.6	-0.2	+0.2	+0.6	+1	Mean	-1	-0.6	-0.2	+0.2	+0.6	+1	
Panel A: All respondents															
Violation	-0.83	0.63	0.32	0.04	0.00	0.00	0.00	-0.77	0.56	0.34	0.09	0.00	0.00	0.01	2.04**
Direct Punishment	0.52	0.01	0.02	0.06	0.23	0.43	0.25	0.51	0.01	0.03	0.04	0.25	0.42	0.24	-0.33
No Help	-0.50	0.24	0.41	0.28	0.04	0.02	0.02	-0.52	0.20	0.45	0.30	0.04	0.00	0.00	0.10
No Help (Violation)	-0.13	0.08	0.24	0.31	0.22	0.09	0.06	-0.20	0.07	0.25	0.35	0.24	0.07	0.01	-1.01
Observations	251							255							
NH vs. NH(V)	-10.11***							-9.74***							
Panel B: Scenario w/ Risk															
Violation	-0.81	0.60	0.35	0.04	0.00	0.00	0.01	-0.74	0.48	0.41	0.10	0.00	0.01	0.01	2.34**
Direct Punishment	0.49	0.02	0.01	0.06	0.26	0.40	0.24	0.48	0.01	0.04	0.05	0.28	0.39	0.23	-0.32
No Help	-0.51	0.24	0.41	0.29	0.03	0.02	0.01	-0.51	0.20	0.45	0.29	0.04	0.01	0.01	0.25
No Help (Violation)	-0.15	0.07	0.25	0.34	0.21	0.08	0.05	-0.21	0.07	0.29	0.35	0.23	0.05	0.02	-0.94
Observations	156							153							
NH vs. NH(V)	-8.32***							-7.24***							
Panel C: Scenario w/o Risk															
Violation	-0.86	0.68	0.28	0.03	0.00	0.00	0.00	-0.82	0.68	0.24	0.08	0.00	0.00	0.01	0.38
Direct Punishment	0.57	0.00	0.02	0.05	0.18	0.46	0.28	0.56	0.02	0.01	0.02	0.22	0.47	0.26	-0.23
No Help	-0.48	0.24	0.40	0.25	0.05	0.03	0.02	-0.53	0.21	0.45	0.30	0.04	0.00	0.00	-0.16
No Help (Violation)	-0.11	0.11	0.23	0.25	0.23	0.11	0.07	-0.18	0.09	0.21	0.35	0.26	0.09	0.00	-0.56
Observations	95							102							
NH vs. NH(V)	-5.81***							-6.51***							

Notes: Social appropriateness of each action for the Child and Alone conditions are reported separately. Options range from very socially inappropriate (-1) to very socially appropriate (+1). Entries denote the share of respondents choosing the option belonging to the corresponding column. The final column displays the absolute value of the test statistics of Wilcoxon Rank-Sum (or, Mann-Whitney U) tests comparing the distributions under both conditions. The rows “NH vs. NH(V)” display the test statistics of Wilcoxon Signed Rank tests examining whether the social appropriateness of not helping is higher following a norm violation. All tests are two-sided. ** $p < 0.05$ *** $p < 0.01$.

this confound.¹⁷ When we split the results according to the two different versions of the vignettes, we obtain the results in Panels B (“Scenario w/ Risk”) and C (“Scenario w/o Risk”). As can be seen, the significant difference is driven entirely by the first version of the vignettes, suggesting that the perceived risk for the child played an important role in the vignette study.¹⁸ In either case, parents did not deem direct punishment more appropriate, which allows us to exclude the possibility that parents in our field study punished more in the presence of a child because they believed this is the social norm. Instead, they did so because they felt more inclined to enforce the same social norm.

Similarly, not providing help to a stranger was deemed equally socially inappropriate in the presence and the absence of a child. We see a slight divergence for not providing help to a violator, but this difference fails to be significant. Finally, note that in both conditions not helping a stranger was deemed significantly less inappropriate when this stranger littered before the helping opportunity presented itself (all respondents: $z_C = -10.113, p < 0.001$; $z_A = -9.742, p < 0.001$; Wilcoxon Signed-Rank tests).

We acknowledge that the respondents to our vignette study are probably not rep-

¹⁷As noted by a reviewer, the littering per se may be perceived as more inappropriate when plastic is involved, whereas a banana peel is organic waste. However, it should not be perceived differently across the conditions (Child or Alone) for other reasons than the violation itself. The problem with the scenario mentioning the banana peel only is that it may be perceived as risky only when the child is around, thereby creating a confound across the two conditions. Moreover, the description with the plastic bag matches the actual scenes more closely.

¹⁸Note that the overall effect is nonetheless quite small and not such as to justify a substantial change in behavior like the one observed in our field experiment. It also ceases to be significant at 5% level if we control for the socio-demographic characteristics of the respondents.

representative of the population of parents in Lyon, because they self-selected. These respondents may be more involved at school or in the education of their child, or care more about social norms. However, if this is the case, we suspect that these parents may be more sensitive to the clue about the presence of the child in the description of the different scenarios. Therefore, not finding significant differences across scenarios with and without the child in this population of respondents suggests that this might be a relatively robust finding. Moreover, we tested whether the parents' responses varied according to socio-demographic variables and the order of scenarios. These variables (age, income, number of children, location, and vignette order) are largely insignificant. There are two exceptions (both significant at the 5% level): higher educated respondents deem punishment less appropriate (but they do not perceive the violation differently) and males deem not helping less inappropriate than females in absence of the violation.

A.5.2 Instructions of the Vignette Study

[Translated from French]

Participant Information Statement

1. What does the study involve? This study involves a very brief questionnaire.
2. Who is carrying out the study? The study is being conducted by professors Fabio Galeotti and Marie Claire Villeval from CNRS and the University of Lyon, and Thijs Brouwer from Tilburg University.
3. How much time will the study take? Answering this questionnaire will take approximately 4 minutes to complete.
4. Can I withdraw from the study? Participating in this questionnaire is completely voluntary. If you do consent, you can withdraw at any time during the questionnaire. Withdrawal from the questionnaire means that you renounce to the chance of winning an electronic tablet, but it will not affect your relationship with the researchers or staff at the CNRS, the University of Lyon or Tilburg University.
5. Will anyone else know the results? All aspects of the questionnaire will be confidential and only the researchers will have access to the responses. A report of the study may be submitted for publication, but all information will only be used in an aggregated form, no personal information will be made public.
6. Will the study benefit me? Responding to the questionnaire will not lead to any payment. However, it will be proposed to the participants to enter a lottery in which one participant will be randomly selected to earn an electronic tablet.
7. Can I tell other people about the study? The researchers request, that for the purpose of maintaining study integrity, you do not share with anybody the nature of the questionnaire.
8. What if I require further information about the study or my involvement in it? If you have specific questions regarding the study, please feel free to contact Marie Claire Villeval by email at villeval@gate.cnrs.fr

Scenarios Below, you will read three short scenarios. In each of the scenarios, you are asked to evaluate the described behavior, choosing between six options ranging from “Very Socially Inappropriate” to “Very Socially Appropriate”. By “socially appropriate” we mean a behavior judged correct and ethical by the majority of people. The objective is to choose, for each scenario, the most common option selected by all other respondents to this questionnaire (all parents with at least one child registered in an elementary school in Lyon).

If you are randomly selected at the end of the study, you will win an electronic tablet (model iPad 32 Go) if your response to one randomly selected question in these scenarios matches the most common response given by all other respondents to the same question. For example, if the most common answer is “Very Socially Inappropriate”, you would receive the tablet if you also answered “Very Socially Inappropriate”. If the most common answer is “Very Socially Appropriate”, you would receive the tablet if you also answered “Very Socially Appropriate”. You will be informed by email if you have won the electronic tablet after all responses have been collected. Please press “Next” to continue.

Vignette 1: Littering + Child /Alone/

A passerby is walking on the street in proximity of an elementary school. This passerby carelessly throws a plastic bag containing food waste on the sidewalk at the sight of a parent with a 6-year-old child / *a parent who has just dropped his/her child at the school* / and no one else around. How would you evaluate the behavior of the passerby? If you give the same response as the majority of the other respondents, you may win a tablet.¹⁹

Please choose one option below:

- Very Socially Inappropriate
- Socially Inappropriate
- Somewhat Socially Inappropriate
- Somewhat Socially Appropriate
- Socially Appropriate
- Very Socially Appropriate

Please press ”Next” to continue.

The parent addresses the passerby and asks this passerby to pick up the plastic bag. How would you evaluate the behavior of the parent? If you give the same response as the majority of the other respondents, you may win a tablet.

Please choose one option below:

¹⁹The first version of the vignette emphasized the banana peel more. The introduction of the vignette read: “A passerby is walking on the street in proximity of an elementary school *while eating a banana*. This passerby carelessly throws *the banana peel* on the sidewalk...” We decided to change this because it did not portray the scene accurately and because the perceived risk of slipping might confound the parents’ perception of the severity of the violation.

- Very Socially Inappropriate
- Socially Inappropriate
- Somewhat Socially Inappropriate
- Somewhat Socially Appropriate
- Socially Appropriate
- Very Socially Appropriate

Please press "Next" to continue.

Vignette 2: Help + Child /Alone/

A passerby is walking on the street in proximity of an elementary school, while carrying a bag containing folders. The passerby accidentally drops all the folders on the ground at the sight of a parent with a 6-year-old child */a parent who has just dropped his/her child from school/* and no one else around. The parent **does not** go to help the passerby with picking up the folders.

How would you evaluate the behavior of the parent? If you give the same response as the majority of the other respondents, you may win a tablet.

Please choose one option below:

- Very Socially Inappropriate
- Socially Inappropriate
- Somewhat Socially Inappropriate
- Somewhat Socially Appropriate
- Socially Appropriate
- Very Socially Appropriate

Please press "Next" to continue.

Vignette 3: Help + Littering + Child /Alone/

A passerby is walking on the street in proximity of an elementary school, while carrying a bag containing folders. This passerby carelessly throws a plastic bag containing food waste on the sidewalk at the sight of a parent with a 6-year-old child */a parent who has just dropped his/her child from the school/* and no one else around. Few instants afterwards, this passerby accidentally drops all his/her folders on the ground. The parent **does not** go to help the passerby with picking up the folders.

How would you evaluate the behavior of the parent? If you give the same response as the majority of the other respondents, you may win a tablet.

Please choose one option below:

- Very Socially Inappropriate
- Socially Inappropriate
- Somewhat Socially Inappropriate
- Somewhat Socially Appropriate
- Socially Appropriate
- Very Socially Appropriate

Please press "Next" to continue.

Before we finish, we would like to ask you a few questions about yourself.

- What is your gender?
 Male Female
- What is your highest educational degree obtained?
 Primary school Less than high school High school diploma or equivalent
 Undergraduate degree Post-graduate degree
- What year were you born (*e.g.*, 1970)?
- How many children do you have?
 0 1 2 3 or more
- What is their gender? How many sons: ____ How many daughters: ____
- What is their age? Your son(s): ____ Your daughter(s): ____
- What is your household monthly earnings category:
 < 2000 Euro 2000-3999 Euro 4000-5999 Euro 6000 Euro and more
- If you live in Lyon, what is your district?
 1 2 3 4 5 6 7 8 9 I don't live in Lyon

Please press "Next" to continue.

Earnings

You may win an electronic tablet if you are randomly selected among all the respondents at the end of our study, and if your response in one randomly selected scenario matches the most common response given by the other respondents. If you are willing to participate in this lottery, please enter your email address below so that we can contact you if you have won the tablet.

Thank you for taking time out of your busy life to participate to this study. If you have any questions concerning this study, you can contact us at villeval@gate.cnrs.fr

A.6 An Online Video Study for Norm Elicitation

In this section, we propose a second robustness test of our assumption that individuals do not consider littering a stronger norm violation in the presence than absence of a child. We first introduce the design of the study. Then, we report on the main findings.

A.6.1 Design

This new survey was conducted in December 2021 and it was not pre-registered. We recruited 243 participants on the Prolific platform in France and Belgium and the survey was conducted in French. We selected these countries to be in a similar normative environment as in our field experiment; we included Belgium (58 respondents, in total) to have enough potential participants. We selected participants above 29 years old to increase the chance of having parents among the respondents.

We preliminary recorded three videos featuring a person violating the littering norm (throwing a can of soda) in front of a bystander in a natural and clean environment. In one video the bystander was walking alone, while in the two other videos the bystander was accompanied either by a child or by another adult (videos are available at: <https://bit.ly/331Yhju>). In the survey, we randomly assigned one of the videos to each respondent. After watching the video, the respondents had to rate the social appropriateness of the violator’s behavior in the video on a six-point scale ranging from “very socially inappropriate” to “very socially appropriate” (see details in subsection A.6.3 below). They were instructed that they could earn 0.5 Euro if they chose the option that was chosen by the majority of the other respondents (that is, participants over 29 years old, living in France or Belgium, and recruited on Prolific).

Then, participants had to imagine that the bystander on the video expressed verbally her disapproval to the violator, and rate the appropriateness level of this reaction, using the same scale and incentives as before.²⁰ After that, we also elicited the personal norms by asking the participants to answer the same two previous questions but according to their own personal opinion and not from the perspective of the majority of the other participants. These two responses were not incentivized. Finally, we asked the participants to indicate the reasons that might cause the bystander to disapprove the violator, and rank them in order of importance. The respondents could indicate one or more (up to four) reasons. A ranking of 1 indicated the most important reason, while a ranking of 4 the least important reason. They could also report that they had no idea. In that case, no ranking could be established. At the end of the study, participants also answered a socio-demographic questionnaire.

A.6.2 Main Findings

Results on the perceived norms are presented in Table A12. As before, we coded the ratings as equidistant values on a range from -1 to 1 , with the former indicating “very (socially) inappropriate” and the latter “very (socially) appropriate”. First, we note that social and personal norms were very similar in all the video conditions, both when

²⁰At the end of the study, one of the two questions was randomly selected for payment.

respondents evaluated the behavior of the violator and the behavior of the bystander.²¹ Second, we find no difference in both personal and social norms across video conditions. The behavior of the violator was perceived as equally inappropriate in all conditions both from the perceived viewpoint of society (Kruskal–Wallis test, $p = 0.703$; $p > 0.1$ for all bivariate comparisons) and from the respondents’ own personal viewpoint (Kruskal–Wallis test, $p = 0.328$; $p > 0.1$ for all bivariate comparisons). Similarly, the respondents deemed the punishing behavior of the bystander equally appropriate in all conditions, both from the perceived viewpoint of society (Kruskal–Wallis test, $p = 0.794$; $p > 0.1$ in all bivariate comparisons) and from their own viewpoint (Kruskal–Wallis test, $p = 0.983$; $p > 0.1$ in all bivariate comparisons).²² These results confirm our findings reported in Appendix A.5, and support the hypothesis that punishment in the presence of a child is not driven by the perception that littering is a more serious violation or that direct punishment is a more appropriate action when a child is present. This is true irrespective of whether the appropriateness of these actions is evaluated from the society’s view or the respondents’ personal convictions.

In the survey, we also asked the respondents to indicate and rank the main reasons that might induce the bystander to disapprove the violator in the video. We were interested in verifying whether the respondents indicate a teaching motive directed to the child as a main reason to punish the norm violator in the Child video condition. Other reasons to punish that we considered were “social image” (i.e. showing a good image of oneself to the child), and “norm enforcement” (i.e. enforcing the rules of good social conduct). We group together any other motives mentioned by the respondents into the category “other reasons”.²³ Table A13 reports the average ranking and the distribution of rankings for the different motives indicated by the respondents in the Child video condition.

First, we note that the respondents almost always (more than 97% of the times) indicated “teaching the child” as a reason that might cause the bystander to punish the norm violator. Furthermore, respondents assigned to the teaching reason a ranking of 1 in almost 1/3 of the cases. In addition, teaching the child was, on average, deemed a more important reason to punish the norm violator than social image (Wilcoxon signed-rank test, $p < 0.001$) and “other reasons” ($p < 0.001$). Its importance was comparable to that of enforcing the norm ($p = 0.256$).²⁴

To have a more complete picture of the ranking patterns, we conducted a rank-

²¹Personal ratings were slightly more negative than social ratings when participants evaluated the behavior of the violator, but only when we pool all the video conditions together (-0.95 vs. -0.93 , Wilcoxon signed-rank test, $p = 0.055$). Personal ratings were also slightly more positive than social ratings when participants evaluated the behavior of the bystander. The difference is statistically significant if we pool all the video conditions together (0.64 vs. 0.58 , $p = 0.017$) and in the Child video condition (0.65 vs. 0.57 , $p = 0.065$). In all the other cases, personal and social norms were statistically indistinguishable ($p > 0.1$ for all comparisons).

²²Qualitatively similar results are obtained from OLS regressions using the appropriateness rating as the dependent variable and the video conditions as the independent variables, and controlling for socio-demographic characteristics (age, income, number of children, Belgian nationality, and size of the residing city). The results of these regressions are available upon request.

²³This category includes, for example, environmental or ecological reasons (60% of the times), and general references to civism (35%).

²⁴In performing all these tests, we gave an arbitrary ranking of 99 when a reason was not ranked. This is to account for the fact that an unranked reason is deemed not important by a respondent. Note that the results are unaffected by the number that we assign to the unranked reasons (as long as this is greater than 4).

Table A12: Appropriateness of Behavior

	Alone video						
	Mean	-1	-0.6	-0.2	0.2	0.6	1
Violation (social)	-0.92	0.83	0.16	0	0	0.01	0
Violation (personal)	-0.94	0.84	0.16	0	0	0	0
Punishment (social)	0.57	0.01	0.01	0.1	0.17	0.34	0.37
Punishment (personal)	0.63	0	0.04	0.06	0.18	0.22	0.5
	Child video						
	Mean	-1	-0.6	-0.2	0.2	0.6	1
Violation (social)	-0.92	0.81	0.18	0.01	0	0	0
Violation (personal)	-0.94	0.86	0.13	0.01	0	0	0
Punishment (social)	0.57	0.01	0.04	0.03	0.14	0.5	0.28
Punishment (personal)	0.65	0	0.05	0.06	0.06	0.36	0.46
	Two-adult video						
	Mean	-1	-0.6	-0.2	0.2	0.6	1
Violation (social)	-0.94	0.86	0.14	0	0	0	0
Violation (personal)	-0.97	0.92	0.08	0	0	0	0
Punishment (social)	0.61	0.01	0.01	0.04	0.14	0.46	0.34
Punishment (personal)	0.65	0.01	0	0.1	0.14	0.24	0.51
	K-W	Alone	Alone	Child			
	test	vs.	vs.	vs.	Child	Two-	Two-
		Child	adult	adult		adult	adult
Violation (social)	.703	.728	.624	.402			
Violation (personal)	.328	.785	.145	.247			
Punishment (social)	.794	.891	.626	.506			
Punishment (personal)	.983	.933	.860	.904			

Notes: Appropriateness of each action according to the perceived social norm (*social*) or the personal view (*personal*) for the Alone, Child and Two-adult conditions are reported separately. Options range from very (socially) inappropriate (-1) to very (socially) appropriate ($+1$). Entries denote the share of respondents choosing the option belonging to the corresponding column. The bottom part of the table displays the p-values of Kruskal Wallis tests (column ‘K-W test’) and Mann-Whitney U tests for each pairwise comparison between conditions (Alone vs. Child, Alone vs. Two-adult and Child vs. Two-adult). All tests are two-sided.

Table A13: Reasons to punish a norm violator (Child video condition)

	Teaching	Social image	Norm enforcement	Other reasons
Ranking 1	0.31	0.09	0.45	0.13
Ranking 2	0.50	0.05	0.38	0.03
Ranking 3	0.15	0.56	0.10	0.05
Ranking 4	0.01	0.15	0.03	0.05
No Ranking	0.03	0.14	0.04	0.74
Avg Ranking	1.87	2.91	1.69	2.10

Notes: The table reports the frequencies of each ranking and the average ranking assigned by the respondents to each reason that might cause the bystander to punish the norm violator in the Child condition. The average ranking is computed from observations with a ranking. A ranking of 1 was assigned to the most important reason, while a ranking of 4 to the least important reason. Respondents were free to rank only one or more reasons. They could also indicate that they had no idea. Avg = average.

ordered logit regression (Beggs et al., 1981; Hausman and Ruud, 1987). For each respondent we have four observations, each corresponding to a ranking (from 1 to 4) assigned to a given reason to punish a norm violator (teaching the child, social image, norm enforcement, other reasons).²⁵ Unranked reasons were given an arbitrarily value of 99 to account for the fact that they were deemed less important than the ranked alternatives by the respondents.²⁶ The dependent variable is the ranking assigned to the different reasons. The explanatory variables are dummies for the different reasons (with teaching taken as the baseline category).²⁷ We employed robust standard errors, adjusted for clustering within respondents. Table A14 reports the parameter estimates of the rank-ordered logit regression. The results should be interpreted in terms of preferences over alternative reasons for punishment.

The estimates of Table A14 indicate that, on average, respondents deemed teaching the child and norm enforcement as equally important for punishing a norm violator. Also, teaching was deemed much more important than social image and other reasons. We can also calculate the probabilities that a given reason was ranked first (bottom part of Table A14). The probability that teaching was ranked first was 39%. This confirms that teaching the child is considered a very important motive to punish the norm violator when the child is present.

²⁵The rank-ordered logit model produces different estimates depending on how the ranking order is presented (i.e. whether a higher number means a more important or a less important alternative). If we reverse the order of the ranking, our results do not change qualitatively.

²⁶We can attach any value to incomplete rankings as long as it is different from 1, 2, 3 or 4.

²⁷We did not include the characteristics of the respondents as predictor variables. This is because these variables do not vary between alternatives, and thus their additive effect cannot affect the rankings. If we run a specification with the inclusion of interaction terms between the dummies for the different reasons and the socio-economic characteristics of the respondents, our main conclusions do not change. The results are available upon request.

Table A14: Importance of the reasons to punish a norm violator (Child video condition)

Reason	β
Teaching the child	<i>Ref.</i>
Social image	-1.43*** (0.20)
Norm enforcement	0.21 (0.18)
Other reasons	-2.68*** (0.38)
<hr/> <i>P</i> <hr/>	
Teaching	0.39
Social image	0.09
Norm enforcement	0.49
Other reasons	0.03

Notes: The top part of the table reports the parameter estimates of a Rank-ordered logit regression. The dependent variable is the ranking assigned to each reason to punish a norm violator. All of the estimates are contrasts with the reference category (teaching). The bottom part of the table reports the probabilities that a given reason was ranked first. These probabilities are computed from the estimates of the Rank-ordered logit regression. Ref. = reference category. ** $p < 0.05$ *** $p < 0.01$.

A.6.3 Instructions of the Video Study

[Translated from French]

Participant Information Statement

1. What does the study involve?
This study involves a very brief questionnaire.
2. Who is carrying out the study?
The study is being conducted by professors Fabio Galeotti and Marie Claire Villeval from CNRS and the University of Lyon.
3. How much time will the study take?
Answering this questionnaire will take approximately four minutes to complete.
4. Can I withdraw from the study?
Participating in this questionnaire is completely voluntary. If you do consent, you can withdraw at any time during the questionnaire. Withdrawal from the questionnaire means that you renounce to the earnings for your participation, but it will not affect your relationship with Prolific or the researchers or staff at the CNRS and the University of Lyon.
5. Will anyone else know the results?
All aspects of the questionnaire will be confidential and only the researchers will have access to the responses. A report of the study may be submitted for publication, but all information will only be used in an aggregated form, no personal information will be made public.
6. Will the study benefit me?
Responding to the questionnaire will pay you 1 Euro plus a bonus that will depend on your answers to the questionnaire.

7. Can I tell other people about the study?

The researchers request, that for the purpose of maintaining study integrity, you do not share with anybody the nature of the questionnaire.

8. What if I require further information about the study or my involvement in it?

If you have specific questions regarding the study, please feel free to contact Marie Claire Villeval by email at villeval(at)gate.cnrs.fr

Please use a computer (and not a cell phone) and an updated version of your browser (and other than Firefox) to complete this study. Throughout the study, please click only once on each button. If you are willing to continue, please check the following box and click on "Next".

I consent to participate.

[Next]

Welcome! Thank you for participating in this study.

Please enter your Prolific ID and click on "Validate" to continue.

Please make sure there are no spaces before or after your Prolific ID.

Your Prolific ID:

[Next]

Questionnaire

You will watch a very short video. Then you will be asked to evaluate two types of behavior, choosing between six options ranging from "very socially inappropriate" to "very socially appropriate". By "socially appropriate" we mean a behavior that is judged correct and ethical by the majority of people.

The objective is to choose the most common option selected by all the other respondents to this questionnaire (people over 29 years old who are registered like you on Prolific in France and Belgium).

You will earn 0.5 Euro if your response to one randomly selected question matches the most common response given by all the other respondents to the same question.

For example, if the most common answer is "very socially inappropriate", you will receive 0.5 Euro if you also answered "very socially inappropriate". If the most common answer is "very socially appropriate", you will receive 0.5 Euro if you also answered "very socially appropriate".

Please press "Next".

[*Next*]

Questionnaire

[*Video*]

The video shows a passerby walking down the street. This passerby casually throws a piece of garbage on the ground.

Please watch the video. You can watch it several times. If you are unable to load the video, the solution is to change your browser by logging back into <https://eugen.gate.cnrs.fr/fgmcv/indexReco.php>.

Once you have watched the video, press OK to see the questions.

[*The respondent cannot proceed without launching the video*]

[*OK*]

Questionnaire

[*Video still available*]

How would you evaluate the behavior of the passerby who littered?

If you give the same response as the majority of the other respondents, you may earn 0.5 Euro.

Please choose one option below:

- Very socially inappropriate
- Socially inappropriate
- Somewhat socially inappropriate
- Somewhat socially appropriate
- Socially appropriate
- Very socially appropriate

[*OK*]

Questionnaire

[*Video still available*]

Imagine that another passerby (the person with the dark green coat) addresses the passerby who littered and asks this passerby to pick up the trash.

How would you evaluate this behavior of the person with the dark green coat?

If you give the same response as the majority of the other respondents, you may earn 0.5 Euro.

Please choose one option below:

- Very socially inappropriate
- Socially inappropriate
- Somewhat socially inappropriate
- Somewhat socially appropriate
- Socially appropriate
- Very socially appropriate

[*OK*]

Questionnaire

[*Video still available*]

We will now ask you to evaluate the same two types of behavior, but this time in your own opinion, independent of the opinions of others.

You will have the opportunity to choose from six options ranging from "very inappropriate" to "very appropriate". By "appropriate" we mean behavior that you feel is correct and ethical. **We are interested in your personal opinion, regardless of the opinion of others.**

We ask that you answer as accurately as possible, giving your own opinion truthfully. There is no right or wrong answer; you will not receive additional compensation for your answers to these questions.

Please press OK.

[OK]

————- **Questionnaire**

[*Video still available*]

How would you evaluate **personally** the behavior of the passerby who littered?
Please choose one option below:

- Very socially inappropriate
- Socially inappropriate
- Somewhat socially inappropriate
- Somewhat socially appropriate
- Socially appropriate
- Very socially appropriate

[OK]

————-

Questionnaire

[*Video still available*]

As before, imagine that another passerby (the person with the dark green coat) addresses the passerby who littered and asks this passerby to pick up the trash.

How would you evaluate **personally** this behavior of the person with the dark green coat?

Please choose one option below:

- Very socially inappropriate
- Socially inappropriate
- Somewhat socially inappropriate
- Somewhat socially appropriate
- Socially appropriate
- Very socially appropriate

[OK]

————-

Questionnaire

What reason(s) do you think might cause the person in the dark green coat to ask the passerby to pick up the trash he has thrown on the ground?

Please rank these reasons in order of importance, reason 1 being the most important reason, and reason 4 the least important reason. You may, of course, indicate one or more reasons. [*Alone condition*]

- Reason 1
- Reason 2
- Reason 3
- Reason 4

[OK]

Please rank these reasons in order of importance by typing 1 below the most important reason and 4 below the least important reason. You may, of course, select one or more reasons. [*Child and Two-Adult conditions*]

- Showing a good image of oneself to one's child. [*Child condition*]
Showing a good image of oneself to the other adult. [*Two-Adult condition*]
- Enforcing the rules of good social conduct.
- Teaching one's child that littering should not go unpunished. [*Child condition*]
Reminding the other adult that accompanies that littering should not go unpunished. [*Two-Adult condition*]
- Other reason
Please specify if you have chosen "Other reason" :
- I have no idea

Questionnaire

Before we finish, we would like to ask you a few questions about yourself.

- What is your gender?
 Male Female Neutral
- What year were you born (*e.g.*, 1970)?
- What is your highest educational degree obtained? [*Due to a glitch in the software, this question was not recorded*]
 No certificate Secondary school certificate Occupational training certificate (CAP, BEP) High school certificate (Baccalaureat) Undergraduate University degree (DUT, Bachelor, ...) Master degree or above

- How many children do you have in your care?
 0 1 2 3 or more
- What is your household monthly net after-tax earnings category:
 < 1500 Euro 1500-2499 Euro 2500-3499 Euro 3500-4499 Euro 4500 Euro and more
- In which size of commune are you living?
 A commune with less than 2000 inhabitants Between 2000 and 9999 Between 10000 and 49999 Between 50000 and 399999 Beyond 400 000 inhabitants
- In which country do you live most of the time?
 Belgium France
- In which region do you live most of the time? [*If Belgium*]
 Brussels capital Flanders Wallonia
- In which region do you live most of the time? [*If France*]
 Auvergne-Rhône-Alpes Bourgogne-Franche-Comté Bretagne Centre-Val de Loire Corse Grand Est Hauts-de-France Ile-de-France Normandie Nouvelle-Aquitaine Occitanie Pays de la Loire Provence-Alpes-Côte d'Azur Guadeloupe Martinique Guyane La Réunion Mayotte

Thank you!

Thank you for taking time to participate in this study.

We will draw one of the first two questions above and for the question drawn, you will earn 0.5 Euro in addition to your fixed payoff of 1 Euro if your answer is the same as the answer of the majority of the other participants.

You will receive your earnings on your Prolific account in a few days.

If you have any questions about this study, you can contact us at [villeval\(at\)gate.cnrs.fr](mailto:villeval(at)gate.cnrs.fr)

Please press the Prolific button to confirm your participation.
Please press "Next" to continue.