Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

PROCEEDINGS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The strategy of psychopathy: primary psychopathic traits predict defection on low-value relationships

Matthew M. Gervais, Michelle Kline, Mara Ludmer, Rachel George and Joseph H. Manson

Proc. R. Soc. B 2013 **280**, 20122773, published 27 February 2013

Supplementary data "Data Supplement" http://rspb.royalsocietypublishing.org/content/suppl/2013/02/25/rspb.2012.2773.DC1.html

References **This article cites 58 articles, 5 of which can be accessed free** http://rspb.royalsocietypublishing.org/content/280/1757/20122773.full.html#ref-list-1

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Proc. R. Soc. B* go to: **http://rspb.royalsocietypublishing.org/subscriptions**

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

PROCEEDINGS OF THE ROYAL SOCIETY B

rspb.royalsocietypublishing.org

Research

Cite this article: Gervais MM, Kline M, Ludmer M, George R, Manson JH. 2013 The strategy of psychopathy: primary psychopathic traits predict defection on low-value relationships. Proc R Soc B 280: 20122773. http://dx.doi.org/10.1098/rspb.2012.2773

Received: 21 November 2012
Accepted: 30 January 2013

Subject Areas:
behaviour, cognition, evolution

Keywords:
cooperation, conversation, evaluation, personality, psychopathy

Author for correspondence:
Matthew M. Gervais
e-mail: mgervais@ucla.edu

Electronic supplementary material is available at http://dx.doi.org/10.1098/rspb.2012.2773 or via http://rspb.royalsocietypublishing.org.

Royal Society Publishing
Informing the science of the future

The strategy of psychopathy: primary psychopathic traits predict defection on low-value relationships

Matthew M. Gervais[1,2], Michelle Kline[1,2], Mara Ludmer[1], Rachel George[1,3] and Joseph H. Manson[1,2]

[1]Department of Anthropology, [2]Center for Behavior, Evolution and Culture, and [3]Center for Language, Interaction, and Culture, University of California, Los Angeles, CA 90095-1553, USA

Recent evidence suggests that psychopathy is a trait continuum. This has unappreciated implications for understanding the selective advantage of psychopathic traits. Although clinical psychopathy is typically construed as a strategy of unconditional defection, subclinical psychopathy may promote strategic conditional defection, broadening the adaptive niche of psychopathy within human societies. To test this, we focus on a ubiquitous real-life source of conditional behaviour: the *expected relational value* of social partners, both in terms of their quality and the likely quantity of future interactions with them. We allow for conversational interaction among participants prior to their playing an unannounced, one-shot prisoner's dilemma game, which fosters naturalistic interpersonal evaluation and conditional behaviour, while controlling punishment and reputation effects. Individuals scoring higher on factor 1 (*callous affect*, *interpersonal manipulation*) of the Levenson self-report psychopathy scale defected conditionally on two kinds of low-value partners: those who interrupted them more during the conversation, and those with whom they failed to discover cues to future interaction. Both interaction effects support the hypothesis that subclinical primary psychopathy potentiates defection on those with low expected relational value. These data clarify the function and form of psychopathic traits, while highlighting adaptive variation in human social strategies.

1. Introduction

Humans are remarkable among animals for their propensity to cooperate across a range of non-kin relationships and social scales [1–5]. Against this backdrop, clinical psychopaths stand out as distinctly inhuman. Characterized by a constellation of antisocial traits and behaviours, including lack of empathy, lack of remorse, grandiosity, impulsivity, manipulation, aggression and social norm violations [6], clinically diagnosed psychopaths fail to sustain long-term reciprocal mutualisms, and they are alarmingly insensitive to reputation concerns and to the threat of punishment [7]. Clinical psychopaths apparently lack the very social emotions hypothesized to sustain relationships [8], and they are less likely to cooperate in social dilemmas [9], one-shot or otherwise.

Despite vulnerability to spoiled reputations and sanctions, psychopathy appears to be a stable genetic accompaniment to human nature across societies, albeit at relatively low frequencies [7]. What explains the persistence of this species-atypical personality trait?

Among the adaptationist theories of psychopathy (reviewed by Glenn *et al.* [10]), most (e.g. [11]) characterize it as a heritable, reliably developing 'defector' type that is selected for either (i) when rare (the *frequency dependence* hypothesis) or (ii) when able to escape sanctions through mobility or anonymous interactions (the *social ecology* hypothesis). These hypotheses are consistent with models of the evolution of cooperation in which unconditional defectors persist at low frequencies under certain conditions (e.g. [12]).

© 2013 The Author(s) Published by the Royal Society. All rights reserved.

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

Empirically, however, psychopathic traits do not usually manifest as a categorical 'defector' type. In fact, the discrete classification of clinical psychopathy captures only one tail of a psychopathic trait distribution that extends well into non-institutionalized populations [13,14]. Few evolutionary approaches have considered the implications of continuous variation in psychopathic traits. This is unfortunate, because the fitness advantages of *subclinical* psychopathy may hold the key to explaining the persistence of clinical psychopathy within human populations [15].

Importantly, both clinical and subclinical psychopathy subsume two factors: *primary*, or factor 1 (F1), psychopathy; and *secondary*, or factor 2 (F2), psychopathy [13,16]. Although correlated [7], these two factors converge with distinct personality constructs [17–20], and are amenable to distinct ontogenetic and evolutionary explanations [10,11,21]. While secondary psychopathy taps impulsiveness and anxiety, primary psychopathy captures callous affect, and converges with Machiavellianism and low honesty/humility—traits at the core of psychopathic antisociality. Primary psychopathy is the focus of both the frequency dependence and social ecology hypotheses.

Belying the 'defector' type portrayal of primary psychopathy, several findings from economic games provide evidence for adaptive behavioural flexibility as a function of subclinical primary psychopathy, even under conditions that apparently disfavour clinical psychopathy. First, those higher in subclinical primary psychopathy may defect in one-shot games where there are no long-term benefits to cooperation [22], while cooperating in repeated interactions where sustained mutual cooperation is profitable [23]. By contrast, clinical psychopaths show a strong inclination to defect even in repeated games, when long-term cooperation would maximize pay-offs [9].

Second, those higher in subclinical psychopathy appear especially sensitive to the prospect of punishment. In an ultimatum game (UG), individuals scoring higher on the *Machiavellian egocentricity* subscale of the Psychopathic Personality Inventory [24] make higher offers when they face punishment for offering too little, while accepting low offers to their own benefit [22]. By contrast, clinical psychopaths do not make higher UG offers than controls, and they are especially willing to reject low offers [25].

Such results are complemented by studies of Machiavellianism, which is highly correlated with primary psychopathy [17] and may afford a similar adaptationist explanation [26]. High Machs are (i) less giving in a dictator game, yet more giving in a UG that includes the possibility of punishment [27,28]; (ii) less giving in the first round of a five-round public goods game [29]; (iii) more likely to accept low offers in a one-shot UG [30]; (iv) less likely to reciprocate trust as player 2 in a one-shot trust game [31]; and (v) more responsive to others' decisions in a multi-round public goods game [29]. Of note, both those higher in psychopathy [23] and those higher in Machiavellianism [27] show patterns of brain activation during social dilemma decisions that are consistent with economically strategic behaviour.

These findings point to a potentially significant source of fitness benefits in subclinical primary psychopathy: the flexible exploitation of others through strategic responsiveness to explicit parameters such as interaction length or sanctions. This implies that the fitness advantages to flat affect and instrumental aggression may not be restricted to conditions that favour 'defector' types, such as rare-type advantage and high mobility.

One potentially ubiquitous context for strategic social behaviour involves the characteristics of interaction partners, or their relational 'affordances', which should critically influence social strategies when exogenous incentives are held constant [32,33]. Low-value social partners are potentially profitable targets of exploitation, given that they have few benefits to withhold in the future. The primary psychopathy continuum may correspond to a relative willingness to strategically exploit or invest in others based upon their appraised utility—which might motivate defection if partners are deemed of low-value, including if future interactions are unlikely, but which should motivate cooperation in interactions with valuable, long-term partners [34]. This stands in stark contrast to simplistic rules such as 'always defect' or 'defect whenever possible', the predominant strategies assumed in the psychopathy literature.

Unfortunately, previous adaptationist theories of psychopathy incorporate only the general characteristics of social environments; they say little about how psychopathy may influence reactions to the traits or behaviours of specific individuals. Likewise, in previous studies of subclinical psychopathy, participants have interacted with software algorithms, confederates, sham co-participants or anonymous, unseen co-participants lacking cues to relational value. In one exception, those higher in psychopathy showed a shift in brain activation towards strategic cooperation only when playing a repeated prisoner's dilemma game (PD) with an in-group member [23], suggesting sensitivity to partner value. Consistent with a complementary sensitivity, those higher in subclinical primary psychopathy are especially perceptive of potential victim vulnerability [35].

The current study focuses on relational value as a source of conditional social behaviour in subclinical primary psychopathy. Specifically, we examine the influence of subclinical primary psychopathy on cooperation following face-to-face interaction. By combining methods from linguistic anthropology, social psychology and experimental economics, we present the first research to code verbal behaviour among strangers in detail, and link this to psychopathic traits and to decisions in an unannounced economic game, a one-shot PD. We hypothesize that, following conversational interaction, and in the absence of explicit punishment or reputation formation, individuals higher in self-reported primary psychopathy will opportunistically exploit those who lack cues to high relational value, while cooperating with potentially high-value partners.

Operationally, we focus on two aspects of *expected relational value*, defined as the product of (i) the net benefits one's partner is likely to provide in a given interaction (*quality*) and (ii) the number of future interactions that are likely (*quantity*). We operationalize this in two predictions.

First, we expect more defection by individuals high in F1 psychopathy towards those who interrupt them more frequently. Interruptions cue low partner quality by indexing pairwise disrespect [36], which is widely inferred from pairwise violations of politeness norms [37], and which predicts received indifference, intolerance and even exploitation [38]. Interruptions reduce liking [36] and generate low ratings of leadership [39], suggesting that interruptions garner status and influence through force-based dominance, not prestige [40]—a costly strategy to have directed at oneself, especially in a putatively egalitarian context.

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

Second, we expect more defection by high F1s towards those with whom they fail to discover 'common ground' [41], or cues to the *quantity* of future interactions; in a college student population, these include sharing an academic major, a residence hall and other indicators of a shared social network (see below). The absence of common ground shrinks the shadow of the future and renders the current interaction truly one-shot, lowering overall expected relational value.

We also expect high F1 defection on low-value partners to occur unilaterally, irrespective of predicted partner behaviour. We include actor predictions of recipient game play to test this.

Finally, we expect a positive relationship between F1 psychopathy and total game-play profit, resulting from a combination of selective defection *by* individuals high in F1, and preferential cooperation *towards* those high in F1 owing to their putative glibness and charm [6].

As covariates previously shown to influence social dilemma decisions, we include measures of co-participant facial attractiveness [42], actor socioeconomic status (SES) [43] and common ethnicity [44], as well as a novel measure of subcultural similarity based on clothing and self-adornment. These last two measures allow us to distinguish our 'common ground' measure of expected future interaction from general similarity.

2. Methods

(a) Participants

This study has two phases involving separate groups of participants: (i) those who engaged in the conversations and the PD (*conversation participants*), and (ii) those who later rated photographs of conversation participants for facial attractiveness and cultural style similarity (*raters*). Conversation participants ($n = 105$, 60 female) were native English-speaking university students with a median age of 19 years, and a diverse ethnic composition: 36 white, 36 Asian, five Latino/a, five African-American, three Pacific Islander, two Middle Eastern, 17 bi-racial and one 'other'. We advertised the study as 'small talk among strangers' with a $10 show-up payment (US dollars), but we did not advertise the potential game-play earnings. Raters ($n = 70$, 49 female) were recruited from the same participant pool as the conversation participants, but during the following academic year. Raters were not asked their age or ethnicity. See the electronic supplementary material for discussion of the raters and rating protocols, as well as additional details regarding the participants, measures, procedure, transcription system and inter-rater reliability.

(b) Measures

Conversation participants played a simultaneous one-shot PD with each of two other participants. In each game, participants had an endowment of $3 to 'transfer' or 'keep'. Participants were told that their co-players each had the same choice, and that transferred money would be doubled, so that if they transferred $3, the recipient would receive $6. Resulting total pay-offs are shown in table 1. Participants were also truthfully informed that one participant in each triad would receive a computer-generated random pay-off in order to obscure players' choices and thereby protect confidentiality [45].

After each PD decision, participants predicted how that co-participant played towards themselves and the third co-participant. Following this, each participant completed the Levenson self-report psychopathy scale (LSRP) [16], a 26-item instrument that produces a two-factor structure (see above),

Table 1. Pay-off matrix for player 1 ('actor') in each prisoner's dilemma game. Pay-offs are in US dollars.

		player 2 ('recipient')	
		cooperate ('transfer')	defect ('keep')
player 1 ('actor')	cooperate ('transfer')	$6	$0
	defect ('keep')	$9	$3

and which is both reliable and externally valid [46]. In the LSRP, participants rate their agreement with statements on a four-point scale, from 'disagree strongly' to 'agree strongly'. An example factor 1 item is 'for me, what's right is whatever I can get away with'; a factor 2 item is 'I am often bored'. Participants were then asked their age, sex, ethnicity and childhood hometown with postcode. Finally, participants were photographed head-to-toe while assuming a relaxed, neutral facial expression.

(c) Procedure

Participants were scheduled in groups of four same-sexed individuals, the last-arriving serving as a 'reserve' in the event that two participants were already acquainted. Participants were kept visually isolated from each other while they read consent/instruction forms, which indicated they would converse with others, that it would be videotaped and that they would be asked some questions after the conversation. Each participant was given a letter-coded name tag ('A', 'B' or 'C') so that they could accurately identify each other during the game-play phase of the procedure. Participants were then brought into an adjoining room. After determining that the conversation participants were strangers to each other, an experimenter instructed them to converse for 10 min on any topics they wished, and then left the room, closing the door. Following the conversation, participants returned to visually separated cubicles containing laptops running z-Tree v. 2.1 [47], which led them through all game play and questionnaire measures. A second experimenter sealed cash pay-offs in envelopes, whereas the first experimenter (ignorant of participants' z-Tree responses) photographed, paid and debriefed each participant separately.

(d) Data analysis

We transcribed and coded the first minute, and two other random segments, of each 10-min conversation. J.H.M. transcribed these selections using a version of the Jefferson transcription system [48], which includes finely detailed recording of the onsets and offsets of talk, gaps in talk and overlap of talk by different speakers. J.H.M. coded these portions for 'interruptions'—overlaps that began at points other than when the speaker was finishing an utterance or turn (a 'transition-relevant' point). J.H.M. also coded the full conversation for discovery of dyadic common ground, which included sharing an academic major, an acquaintance, a residence hall, a hobby or other indications of possible shared social networks. All transcription and coding was carried out while blind to participants' PD decisions and questionnaire responses.

To create facial attractiveness and style similarity variables, we calculated the mean rating across raters, for each conversation participant (attractiveness) or for each dyad (similarity). We standardized the following variables as *z*-scores: individual facial attractiveness (separately for each sex); individual total, F1 and F2 LSRP scores; and dyadic cultural style similarity. As an estimate

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

Table 2. Bivariate logistic regressions of hypothesized independent variables on probability of actor (i) predicting cooperation by recipient and (ii) choosing cooperation towards recipient. $n = 206$ for all tests except income on predictions ($n = 200$), income on game play ($n = 198$), and F1 LSRP on predictions ($n = 208$). Robust s.e. calculated based on number of individual actors (one-half of n).

	(i) bivariate regressions on actor *predictions*		(ii) bivariate regressions on actor *game play*	
independent variable	odds ratio ± s.e.	*p*	odds ratio ± s.e.	*p*
median household income in actor's childhood postcode	1.02 ± 0.01	0.041	1.03 ± 0.01	0.003
recipient's facial attractiveness (standardized by sex)	1.36 ± 0.21	0.047	1.42 ± 0.25	0.049
interruptions min^{-1} of actor by recipient	0.64 ± 0.20	0.161	0.41 ± 0.14	0.010
actor's F1 LSRP score (standardized)	0.74 ± 0.14	0.100	0.68 ± 0.15	0.073
actor's F2 LSRP score (standardized)	0.87 ± 0.15	0.408	0.80 ± 0.16	0.258
common ground (yes = 1)	1.51 ± 0.46	0.175	1.33 ± 0.42	0.336
ethnic similarity (same = 1)	0.99 ± 0.32	0.972	0.96 ± 0.36	0.915
style similarity (standardized)	1.01 ± 0.15	0.922	0.99 ± 0.15	0.957

of SES, we used median income of the childhood postcode participants provided ($n = 99$), based on the US Census Bureau's 2000 database [49]. For each dyad, we calculated two unidirectional interruption rates (A interrupts B, and vice versa) as the number of interruptions per minute of transcribed conversation.

We used bivariate and multivariate logistic regression models that accounted for the non-independence of each person's two game-play choices through the inclusion of robust standard errors of the odds ratios, clustered by individual [50,51]. Total PD profit was calculated as the sum of each participant's earned pay-offs vis-à-vis each of his/her co-participants (range: $0–18). We used two-tailed tests for all bivariate regressions on predictions and game play, and for the multivariate predictive models of game play, but a one-tailed test of the relationship between LSRP scores and total PD profit. The data for this study can be accessed in the Dryad repository [52].

3. Results

All measures showed adequate reliability (see electronic supplementary material). LSRP F1 and F2 were correlated at $r = 0.38$ ($n = 105$, $p < 0.0001$).

Each of 105 participants were asked to play two PDs, yielding an expected sample size of 210 PD choices. Two participants declined to play, reducing the number of transactions to 206. Of these, 136 choices (66%) were to cooperate and 70 (34%) were to defect. Of the 103 participants who made PD choices, 64 (62.1%) chose cooperate towards both co-participants, 31 (30.1%) chose defect towards both co-participants, and eight (7.8%) cooperated towards one co-participant and defected towards the other. Actors' predictions and game play were highly correlated ($\chi^2 = 109.840$, $p < 0.001$).

PD choice was not associated with (i) participant sex, (ii) participant recruitment source (posted flyer or course participation pool), (iii) laptop cubicle position, (iv) assigned letter code or (v) whether actor was the participant facing the camera.

Table 2 shows two sets of bivariate logistic regressions: the effects of the independent variables on the probability of actor predicting cooperate from recipient (column 1), and the effects of those independent variables on the probability of actor choosing to cooperate towards recipient (column 2). Actors were more likely to cooperate when (i) actor's childhood postcode had a higher median household income, (ii) recipient was more facially attractive and (iii) recipient interrupted actor less frequently. Actors with lower F1 LSRP scores were marginally more likely to cooperate. Actor's hometown income and recipient's attractiveness were both positively related to actor's predictions of recipient cooperation, whereas recipient's interruptions and actor's F1 LSRP were unrelated to actor's predictions. Actor's F2 psychopathy, discovery of common ground, ethnic similarity and style similarity were not associated with actor predictions or game play.

Table 3 shows two multivariate logistic regression models of the probability of actor choosing to cooperate. Model 1 includes five of the independent variables from table 2, excluding actor's F2 psychopathy, ethnic similarity and style similarity. This model also includes the predicted interactions between actor's F1 LSRP score and (i) recipient interruptions, and (ii) discovery of common ground. This model produced the lowest Akaike information criterion (AIC) [53] score of any model incorporating combinations of these variables, outperforming models using total LSRP or F2 LSRP in place of F1 LSRP, as well as those including ethnic or style similarity in place of common ground. Moreover, none of these excluded variables produced significant main effects or interactions. Model 2 adds actor predictions of recipient game play to model 1.

Holding the other independent variables constant, two main effects are evident in model 1. Actor is more likely to cooperate when (i) recipient is more facially attractive and (ii) actor's childhood postcode had a higher median household income. There are no main effects of F1 LSRP score, interruptions or common ground. However, both interaction effects involving F1 psychopathy are significant. First, when participants fail to discover common ground, those higher on F1 are more likely to defect (figure 1). When participants do discover common ground, F1 LSRP score is not associated with probability of cooperation. Second, participants with higher F1 LSRP scores are more likely to respond to interruptions by defecting (figure 2).

Model 2 adds actor predictions as a covariate. As evident in table 3, predictions have a massive effect on actor game play, and multiply the explanatory power of the model.

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

Table 3. Multivariate logistic regressions of hypothesized independent variables, including two interaction effects, on the probability of actor choosing cooperation ($n = 196$). Shown without (model 1) and with (model 2) actor predictions as a covariate. Robust s.e. calculated based on number of individual actors ($n = 99$). For model 1, Wald $\chi^2 = 28.51$, $p = 0.0002$, pseudo $r^2 = 0.16$. For model 2, Wald $\chi^2 = 54.5$, $p = 0.0000$, pseudo $r^2 = 0.56$.

	multivariate model 1 with two interactions		multivariate model 2 adding actor *predictions*	
AIC score	228.7		134.7	
independent variable	odds ratio ± s.e.	*p*	odds ratio ± s.e.	*p*
actor's prediction of recipient's PD decision towards actor	—	—	68.68 ± 45.61	<0.001
median household income in actor's childhood postcode	1.03 ± 0.01	0.002	1.02 ± 0.01	0.049
recipient's facial attractiveness (standardized by sex)	1.48 ± 0.27	0.032	1.33 ± 0.34	0.261
interruptions min^{-1} of actor by recipient	0.47 ± 0.21	0.094	0.38 ± 0.22	0.088
actor's F1 LSRP score (standardized)	0.95 ± 0.27	0.852	1.14 ± 0.44	0.733
actor/recipient common ground (yes = 1)	1.28 ± 0.49	0.518	0.92 ± 0.38	0.846
interruptions min^{-1} of actor by recipient × actor's F1 LSRP score	0.31 ± 0.14	0.010	0.25 ± 0.12	0.004
actor/recipient common ground × actor's F1 LSRP score	2.40 ± 0.87	0.016	2.70 ± 1.50	0.075

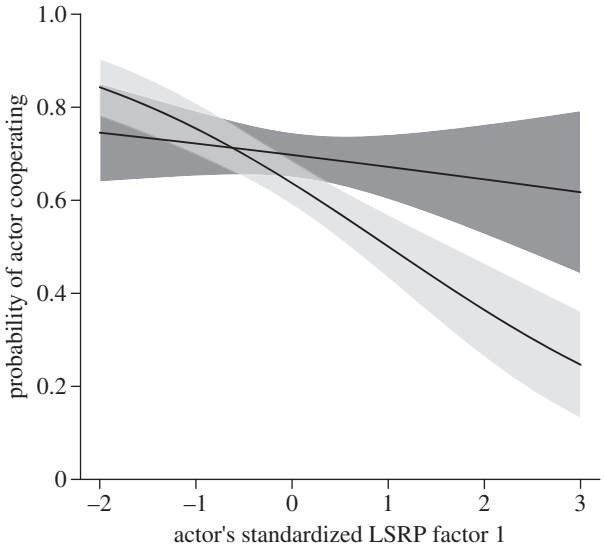

Figure 1. Interaction of actor/recipient 'common ground' and actor's F1 LSRP score, on actor's probability of cooperating with recipient. Actors higher in F1 psychopathy were less likely to cooperate in the absence of found common ground. Shaded regions show ± s.e. Dark grey shading denotes the presence of common ground; light grey shading denotes no common ground.

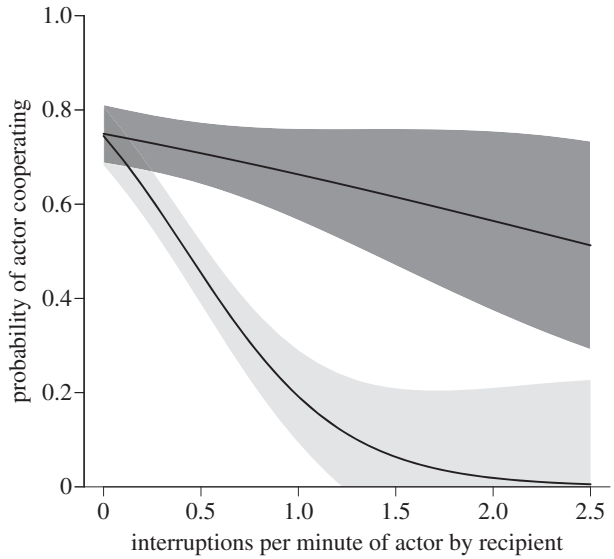

Figure 2. Interaction of recipient interruptions of actor, and actor's F1 LSRP score, on actor's probability of cooperating with recipient. Actors higher in F1 psychopathy were less likely to cooperate the more they were interrupted by recipient. Plot compares the lowest (first; dark grey shade) and highest (fourth; light grey shade) quartile scores on the LSRP factor 1. Shaded regions show ± s.e.

Adding predictions in model 2 has several statistical effects: (i) the main effect of recipient attractiveness disappears, (ii) the main effect of recipient interruptions moves towards marginal significance, and (iii) the interaction of common ground and F1 psychopathy moves away from significance. The interaction of F1 psychopathy and interruptions is strengthened by the addition of actor predictions.

Individuals higher in total LSRP earned higher total profits than those lower in total LSRP ($n = 105$, $p = 0.017$, $\beta \pm$ s.e. = 0.954 ± 0.441). Unexpectedly, LSRP F1 was only marginally associated with total profit ($n = 105$, $p = 0.056$, $\beta \pm$ s.e. = 0.718 ± 0.446), whereas LSRP F2 was significantly associated with total profit ($n = 105$, $p = 0.014$, $\beta \pm$ s.e. = 0.982 ± 0.440). Receipt of cooperation was positively associated with LSRP F2 (clustered logistic regression, $n = 206$, $p = 0.028$, odds ratio = 1.33), but not with LSRP F1 ($p = 0.42$, odds ratio = 1.04) or total LSRP ($p = 0.21$, odds ratio = 1.16).

4. Discussion

By quantifying subtle aspects of open conversations preceding an unannounced one-shot PD, we documented a previously overlooked yet potentially significant source of adaptive benefits in primary psychopathy: defection on low-value social relationships. We found support for two forms of conditional defection as a function of subclinical primary psychopathy: participants scoring higher on F1 of the LSRP were more likely to defect towards co-participants

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

who (i) had interrupted them more frequently, and with whom (ii) they had failed to find common ground. Both variables should factor into a computation of expected relational value, in that they capture, respectively, the quality [34] and the duration [1] of a prospective social relationship.

Defection towards frequent interrupters was not driven by predictions about their PD play, suggesting that this defection was ostensibly unilateral. We also obtained no evidence that such defection was impulsive or driven by anger, nor that higher primary psychopathy potentiated greater frustration, or less self-control, given the same frustrating event [54]; interrupted high-primary participants gave no indication of impulsively responding during the conversation, and were no more likely to interrupt after being interrupted (M.M.G., M.K. & J.H.M. 2013, unpublished data). During the conversation, participants were ignorant of the upcoming game, so they probably were not waiting for confidential revenge. Instead, their reactions to being interrupted took the form of cold defection towards an undesirable social partner, specifically one who had failed to respect them [36]—a slight that suggests a strategy of dominance and cost imposition, and to which people higher in primary psychopathy may be particularly sensitive [55,56]. However, reinforcing our interpretation of this effect as resulting from low partner value and not from a psychopathy-specific problem with authority, neither clinical psychopaths nor those higher in subclinical psychopathy show any deficits in authority-based moral reasoning [57,58].

We also found that a lack of common ground between participants reduced cooperation rates among those higher in primary psychopathy. Common ground included a shared social network, shared course of study, similar residence location or activity patterns and other cues to likely future interaction (see electronic supplementary material). This variable went beyond mere 'personal' conversation [59] and measured discovered aspects of a shared social environment. We found no effects of ethnic or style similarity, suggesting that common ground did not tap mere similarity or compatibility (*sensu* [60]). Thus, the established enhancement effect of face-to-face interaction on cooperation in one-shot social dilemmas [61]—which may depend on cues to future interaction [4,62], and may function through elevated other-regarding preferences [59]—appears qualified among those higher in psychopathy, requiring especially reliable and concrete information about future interaction. Otherwise, those higher in primary psychopathy appear unmoved by face-to-face interaction.

Interestingly, the inclusion of actor predictions in model 2 had divergent effects on our two interaction variables—the interruption effect was strengthened, whereas the common ground effect became only marginally significant. This pattern supports our prediction that interrupters were defected upon unilaterally as low-value prospective partners. However, interpreting the effect of predictions on the common ground interaction is less straightforward. While this effect is consistent with predictions mediating high F1 defection in the absence of common ground, it is also consistent with common ground being a third variable that influences both predictions and game play in parallel [63]—common ground is symmetrical within a dyad in a way that interruptions are not, and perspective taking would lead defectors to predict defection in symmetrical cases even if defection did not hinge on this prediction. Of note, we elicited predictions after game-play decisions, giving the latter causal primacy.

Our findings raise a number of important proximate questions. It remains unknown whether those higher in primary subclinical psychopathy are (i) more likely to devalue others (or devalue others to a greater extent), or (ii) more likely to act antisocially on the same devaluation, perhaps due to less concern with sanctions. We expect both factors to play a role. Future research should investigate the interplay of devaluation and sanctions, such as whether those higher in subclinical primary psychopathy can be encouraged by punishment or reputation to cooperate even with low-value partners, and whether high-primary-psychopathy participants show normal emotional commitment to high-value partners. Of note, those higher in psychopathy can be induced to 'show concern' for others through an in-group manipulation [64], suggesting that high-value relationships can motivate commitment-like behaviour, although not necessarily commitment emotions [8]. Future research could also address the context-specificity of our measures of relational value, specifically interruptions. While interruptions may cue social costs when they occur in putatively egalitarian, non-competitive contexts such as our small talk paradigm, the dominance of an interrupter may index future benefits in the context of coalitionary competition or asymmetrical coordination games [65]. If interruptions are about partner value, we expect their effect on defection to be attenuated in the latter contexts; if, alternatively, those higher in psychopathy are generally put off by interruptions—perhaps owing to their own sense of grandiosity [6,56]—then context may not matter. Finally, future research using a finer-grained measure of individual differences [66] should seek to tease apart the relative contributions to strategic social behaviour of primary psychopathy and Machiavellianism.

In the full model including covariates, two main effects emerged. First, more attractive participants received more cooperation, although not when controlling for predictions of their behaviour. This result accords with previous research showing that more attractive participants receive more cooperation because they are expected to cooperate [42]. As an empirical fact, most humans cooperate flexibly and conditionally [67], and these data underscore that point.

Second, those higher in SES were more likely to cooperate, although this effect was small. This result is opposite to those of Piff *et al.* [43,68], who have found across a range of studies that higher-SES individuals are less generous and more unethical. However, unlike the impersonal contexts studied by Piff *et al.*, our methodology includes face-to-face conversation, which may have brought high-SES participants closer to their interlocutors, making them more likely to share their greater resources. Of note, our sample showed a truncated SES distribution; only 25 per cent of the participants were from postcodes with median incomes less than the nationwide median of $50 000, and only 10 per cent were from postcodes with median less than $40 000. Future research should explore the interaction of SES and social closeness on generosity across a range of SES disparities.

Finally, we found a modest positive effect of LSRP scores on total PD profit. This result emerged from the confluence of two non-significant trends: high LSRP scorers cooperated less, but received more cooperation, than low LSRP scorers. Surprisingly, this result was driven by the receipt of cooperation by those higher in secondary (F2) psychopathy rather than in primary (F1) psychopathy. However, we did not quantify the visible traits or behaviours that elicited such cooperation. Future research should seek to replicate this result and determine its

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

eliciting conditions, for instance by measuring emotional expressivity [69] in secondary psychopathy.

5. Conclusion

The two forms of conditional defection associated with subclinical primary psychopathy and the higher profits of those higher on the LSRP are consistent with the hypothesis that subclinical primary psychopathy is an adaptive social strategy disposed to defect [11]. Our effects further highlight the situational and strategic nature of this defection. This fits with previous research documenting strategic flexibility as a function of subclinical primary psychopathy [22] and the related construct Machiavellianism [29], and builds on it by using more naturalistic methods that highlight partner value as a strategic consideration (see also [70]). Significantly, we documented no costs to subclinical primary psychopathy—they were no more likely to be predicted to defect, and no more likely to be defected upon [71]—raising the possibility that the costs to primary psychopathy are more likely to accrue to those with extreme versions of the trait (e.g. [72]). Although clinical psychopathy may prove adaptive only in social environments that facilitate the unconditional exploitation of others, subclinical psychopathy may function more broadly in environments that present occasional opportunities for defection vis-à-vis particular low-value relationships. When cued to the possibility of a long-term, profitable relationship, subclinical psychopathy appears compatible with the establishment of a cooperative mutualism, however strategic its basis. Primary psychopathic traits may thus persist in a population without selection for extreme (i.e. clinical) psychopathy.

All procedures used in this study were approved by the UCLA Institutional Review Board. We thank Solange Bohling, Erin Campbell, Natalie Chiengkul, Joe Creason, Kimberly Gon, Erika Hoyle, Kotrina Kajokaite, Alexander Kandel, Sabrina Morsello, Kayla Noonen, Sunaina Rao, Sepehr Shakib, Julia Stone and Andrew West for research assistance. Xiao Chen, Rose Medeiros and Phil Ender of UCLA Statistical Consulting provided statistical advice, and Greg Bryant provided logistical support. Cristina Moya shared her R scripts for making figures 1 and 2. This research was supported by an International Society for Human Ethology grant to M.M.G., and a UCLA Faculty Research grant to J.H.M.

References

1. Trivers RL. 1971 The evolution of reciprocal altruism. *Q. Rev. Biol.* **46**, 35–57. (doi:10.1086/406755)
2. Alexander RD. 1987 *The biology of moral systems*. New York, NY: Aldine de Gruyter.
3. Fehr E, Fischbacher U, Gächter S. 2002 Strong reciprocity, human cooperation, and the enforcement of social norms. *Hum. Nat.* **13**, 1–25. (doi:10.1007/s12110-002-1012-7)
4. Delton AW, Krasnow M, Cosmides L, Tooby J. 2011 Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proc. Natl Acad. Sci. USA* **108**, 13 335–13 340. (doi:10.1073/pnas.1102131108)
5. Fehr E, Heinrich J. 2003 Is strong reciprocity a maladaptation? In *Genetic and cultural evolution of cooperation* (ed. P Hammerstein), pp. 55–82. Cambridge, MA: MIT Press.
6. Cleckley H. 1988 (1941) *The mask of sanity*. St Louis, MO: CV Mosby.
7. Hare RD. 2003 *Hare psychopathy checklist-revised (PCL-R)*. Toronto, Canada: Multi-Health Systems Inc.
8. Fessler DMT, Haley KJ. 2003 The strategy of affect: emotions in human cooperation. In *Genetic and cultural evolution of cooperation* (ed. P Hammerstein), pp. 7–36. Cambridge, MA: MIT Press.
9. Mokros A, Menner B, Eisenbarth H, Alpers GW, Lange KW, Osterheider M. 2008 Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *J. Abnorm. Psychol.* **117**, 406–413. (doi:10.1037/0021-843X.117.2.406)
10. Glenn AL, Kurzban R, Raine A. 2011 Evolutionary theory and psychopathy. *Aggress. Violent Behav.* **16**, 371–380. (doi:10.1016/j.avb.2011.03.009)
11. Mealey L. 1995 The sociobiology of sociopathy: an integrated evolutionary model. *Behav. Brain Sci.* **18**, 523–599. (doi:10.1017/S0140525X00039595)
12. Nowak MA, Sigmund K. 2005 Evolution of indirect reciprocity. *Nature* **437**, 1291–1298. (doi:10.1038/nature04131)
13. Hare RD, Neumann CS. 2006 The PCL-R assessment of psychopathy: development, structural properties, and new directions. In *Handbook of psychopathy* (ed. CJ Patrick), pp. 58–90. New York, NY: Guilford.
14. Edens JF, Marcus DK, Lilienfeld SO, Poythress NG. 2006 Psychopathic, not psychopath: taxometric evidence for the dimensional structure of psychopathy. *J. Abnorm. Psychol.* **115**, 131–144. (doi:10.1037/0021-843X.115.1.131)
15. Nettle D. 2005 An evolutionary approach to the extraversion continuum. *Evol. Hum. Behav.* **26**, 363–373. (doi:10.1016/j.evolhumbehav.2004.12.004)
16. Levenson MR, Kiehl KA, Fitzpatrick CM. 1995 Assessing psychopathic attributes in a noninstitutionalized population. *J. Pers. Soc. Psychol.* **68**, 151–158. (doi:10.1037/0022-3514.68.1.151)
17. Lee K, Ashton MC. 2005 Psychopathy, Machiavellianism, and narcissism in the five-factor model and the HEXACO model of personality structure. *Pers. Indiv. Differ.* **38**, 1571–1582. (doi:10.1016/j.paid.2004.09.016)
18. Paulhus DL, Williams KM. 2002 The dark triad of personality: narcissism, Machiavellianism, and psychopathy. *J. Res. Pers.* **36**, 556–563. (doi:10.1016/S0092-6566(02)00505-6)
19. Cima M, Raine A. 2009 Distinct characteristics of psychopathy relate to different subtypes of aggression. *Pers. Indiv. Differ.* **47**, 835–840. (doi:10.1016/j.paid.2009.06.031)
20. Jakobwitz S, Egan V. 2006 The dark triad and normal personality traits. *Pers. Indiv. Differ.* **40**, 331–339. (doi:10.1016/j.paid.2005.07.006)
21. Viding E, Blair RJR, Moffitt TE, Plomin R. 2005 Evidence for substantial genetic risk for psychopathy in 7-year-olds. *J. Child Psychol. Psychiatry* **46**, 592–597. (doi:10.1111/j.1469-7610.2004.00393.x)
22. Curry O, Chester MJ, Viding E. 2011 The psychopath's dilemma: the effects of psychopathic personality traits in one-shot games. *Pers. Indiv. Differ.* **50**, 804–809. (doi:10.1016/j.paid.2010.12.036)
23. Rilling JK, Glenn AL, Jairam MR, Pagnoni G, Goldsmith DR, Elfenbein HA, Lilienfeld SO. 2007 Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry* **61**, 1260–1271. (doi:10.106/j.biopsych.2006.07.021)
24. Lilienfeld SO, Andrews B. 1996 Development and preliminary validation of a self-report measure of psychopathic personality traits in noncriminal populations. *J. Pers. Assess.* **66**, 488–524. (doi:10.1207/s15327752jpa6603_3)
25. Koenigs M, Kruepke M, Newman JP. 2010 Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. *Neuropsychologia* **48**, 2198–2204. (doi:10.1016/j.neuropsychologia.2010.04.012)
26. Wilson DS, Near D, Miller RR. 1996 Machiavellianism: a synthesis of the evolutionary and psychological literatures. *Psychol. Bull.* **119**, 285–299. (doi:10.1037/0033-2909.119.2.285)
27. Spitzer M, Fischbacher U, Herrnberger B, Grön G, Fehr E. 2007 The neural signature of social norm compliance. *Neuron* **56**, 185–196. (doi:10.1016/j.neuron.2007.09.011)
28. Hilbig BE, Zettler I, Heydasch T. 2012 Personality, punishment and public goods: strategic shift towards cooperation as a matter of dispositional honesty-humility. *Eur. J. Personality* **26**, 245–254. (doi:10.1002/per.830)
29. Czibor A, Bereczkei T. 2012 Machiavellian people's success results from monitoring their patterns. *Pers.*

Downloaded from rspb.royalsocietypublishing.org on February 27, 2013

Indiv. Differ. **53**, 202–206. (doi:10.1016/j.paid.2012.03.005)

30. Meyer H-D. 1992 Norms and self-interest in ultimatum bargaining: the prince's prudence. *J. Econ. Psychol.* **13**, 215–232. (doi:10.1016/0167-4870(92)90031-2)
31. Gunnthorsdottir A, McCabe K, Smith V. 2002 Using the Machiavellianism instrument to predict trustworthiness in a bargaining game. *J. Econ. Psychol.* **23**, 49–66. (doi:10.1016/S0167-4870(01) 00067-8)
32. Neuberg SL, Cottrell CA. 2008 Managing the threats and opportunities afforded by human sociality. *Group Dynamics Theor. Res.* **12**, 63–72. (doi:10.1037/1089-2699.12.1.63)
33. Tooby J, Cosmides L. 1996 Friendship and the banker's paradox: other pathways to the evolution of adaptations for altruism. *P. Brit. Acad.* **88**, 119–143.
34. Hruschka DJ, Henrich J. 2006 Friendship, cliquishness, and the emergence of cooperation. *J. Theor. Biol.* **239**, 1–15. (doi:10.1016/j.jtbi.2005.07.006)
35. Wheeler S, Book A, Costello K. 2009 Psychopathic traits and perceptions of victim vulnerability. *Crim. Justice Behav.* **36**, 635–648. (doi:10.1177/0093854809333958)
36. Farley SD. 2008 Attaining status at the expense of likeability: pilfering power through conversational interruption. *J. Noverbal Behav.* **32**, 241–260. (doi:10.1007/s10919-008-0054-x)
37. Brown P, Levinson SC. 1987 *Politeness: some universals in language usage*. Cambridge, UK: Cambridge University Press.
38. Fessler DMT, Gervais MM. 2010 From whence the captains of our lives: ultimate and phylogenetic perspectives on emotions in humans and other animals. In *Mind the gap: the origins of human universals* (eds P Kappeler, J Silk), pp. 261–282. New York, NY: Springer.
39. Karakowski L, McBey K, Miller DL. 2004 Gender, perceived competence, and power displays: examining verbal interruptions in a group context. *Small Group Res.* **35**, 407–439. (doi:10.1177/1046496404263728)
40. Henrich J, Gil-White F. 2001 The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evol. Hum. Behav.* **22**, 165–196. (doi:10.1016/S1090-5138(00)00071-4)
41. Enfield NJ. 2006 Social consequences of common ground. In *Roots of human sociality: culture, cognition and interaction* (eds NJ Enfield, SC Levinson), pp. 399–430. Oxford, UK: Berg.
42. Mulford M, Orbell J, Shatto C, Stockard J. 1998 Physical attractiveness, opportunity, and success in everyday exchange. *Am. J. Sociol.* **103**, 1565–1592. (doi:10.1086/231401)
43. Piff PK, Kraus MW, Côté S, Cheng BH. 2010 Having less, giving more: the influence of social class on prosocial behavior. *J. Pers. Soc. Psychol.* **99**, 771–784. (doi:10.1037/a0020092)
44. Fershtman C, Gneezy U. 2001 Discrimination in a segmented society: an experimental approach. *Q. J. Econ.* **116**, 351–377. (doi:10.1162/003355301556338)
45. Frank RH, Gilovich T, Regan DT. 1993 The evolution of one-shot cooperation: an experiment. *Ethol. Sociobiol.* **14**, 247–256. (doi:10.1016/0162-3095(93)90020-I)
46. Lynam DR, Whiteside S, Jones S. 1999 Self-reported psychopathy: a validation study. *J. Pers. Assess.* **73**, 110–132. (doi:10.1207/S15327752JPA730108)
47. Fischbacher U. 2007 z-Tree: Zurich toolbox for ready-made economic experiments. *Exp. Econ.* **10**, 171–178. (doi:10.1007/s10683-006-9159-4)
48. Jefferson G. 2004 Glossary of transcript symbols with an introduction. In *Conversation analysis: studies from the first generation* (ed. GH Lerner), pp. 13–31. Amsterdam, The Netherlands: John Benjamins.
49. U.S. Census Bureau 2000 American FactFinder. See http://factfinder.census.gov/home/saff/main.html. (accessed 11 January 2010).
50. Rogers WH. 1993 Regression standard errors in clustered samples. *Stata Tech. Bull.* **13**, 19–23.
51. Williams RL. 2000 A note on robust variance estimation for cluster-correlated data. *Biometrics* **56**, 645–646. (doi:10.1111/j.0006-341X.2000.00645.x)
52. Gervais MM, Kline MA, Ludmer M, George R, Manson JH. 2013 Data from: the strategy of psychopathy: primary psychopathic traits predict defection on low-value relationships. *Dryad Digital Repository* (doi:10.5061/dryad.848cc)
53. Akaike H. 1974 A new look at the statistical model identification. *IEEE Trans. Automat. Control* **19**, 716–723. (doi:10.1109/TAC.1974.1100705)
54. Harenski CL, Kiehl KA. 2010 Reactive aggression in psychopathy and the role of frustration: susceptibility, experience, and control. *Br. J. Psychol.* **101**, 401–406. (doi:10.1348/000712609X471067)
55. Glenn AL, Raine A. 2009 Psychopathy and instrumental aggression: evolutionary, neurobiological, and legal perspectives. *Int. J. Law Psychiatry* **32**, 253–258. (doi:10.1016/j.ijlp.2009.04.002)
56. Morrison D, Gilbert P. 2001 Social rank, shame and anger in primary and secondary psychopaths. *J. Forensic Psychiatr.* **12**, 330–356. (doi:10.1080/09585180110056867)
57. Aharoni E, Antonenko O, Kiehl KA. 2011 Disparities in the moral intuitions of criminal offenders: the role of psychopathy. *J. Res. Pers.* **45**, 322–327. (doi:10.1016/j.jrp.2011.02.005)
58. Glenn AG, Iyer R, Graham J, Koleva S, Haidt J. 2009 Are all types of morality compromised in psychopathy? *J. Pers. Dis.* **23**, 384–398. (doi:10.1521/pedi.2009.23.4.384)
59. Buchan NR, Johnson EJ, Croson RTA. 2006 Let's get personal: an international examination of the influence of communication, culture and social distance on other regarding preferences. *J. Econ. Behav. Organ.* **60**, 373–398. (doi:10.1016/j.jebo.2004.03.017)
60. McElreath R, Boyd R, Richerson PJ. 2003 Shared norms and the evolution of ethnic markers. *Curr. Anthropol.* **44**, 122–129. (doi:10.1086/345689)
61. Sally D. 1995 Conversation and cooperation in social dilemmas: a metaanalysis of experiments from 1958 to 1992. *Ration. Soc.* **7**, 58–92. (doi:10.1177/1043463195007001004)
62. Burnham TC, Johnson DDP. 2005 The biological and evolutionary logic of human cooperation. *Analyse and Kritik* **27**, 113–135.
63. MacKinnon DP, Fairchild AJ, Fritz MS. 2007 Mediation analysis. *Annu. Rev. Psychol.* **58**, 593–614. (doi:10.1146/annurev.psych.58.110405.085542)
64. Arbuckle NL, Cunningham WA. 2012 Understanding everyday psychopathy: shared group identity leads to increased concern for others among undergraduates higher in psychopathy. *Soc. Cogn.* **30**, 564–583. (doi:10.1521/soco.2012.30.5.564)
65. King AJ, Johnson DDP, Van Vugt M. 2009 The origins and evolution of leadership. *Curr. Biol.* **19**, R911–R916. (doi:10.1016/j.cub.2009.07.027)
66. Jones DN, Paulhus DL. In press. Introducing the short dark triad (SD3): a brief measure of dark personalities. *Assessment.*
67. Fehr E, Fischbacher U. 2003 The nature of human altruism. *Nature* **425**, 785–791. (doi:10.1038/nature02043)
68. Piff PK, Stancato DM, Côté S, Mendoza-Denton R, Keltner D. 2012 Higher social class predicts increased unethical behavior. *Proc. Natl Acad. Sci. USA* **109**, 4086–4091. (doi:10.1073/pnas.1118373109)
69. Schug J, Matsumoto D, Horita Y, Yamagishi T, Bonnet K. 2010 Emotional expressivity as a signal of cooperation. *Evol. Hum. Behav.* **31**, 87–94. (doi:10.1016/j.evolhumbehav.2009.09.006)
70. Krupp DB, Sewall LA, Lalumière ML, Sheriff C, Harris GT. 2012 Nepotistic patterns of violent psychopathy: Evidence for adaptation? *Front. Psychol.* **3**, 1–8. (doi:10.3389/fpsyg.2012.00305)
71. Manson JH, Gervais MM, Fessler DMT, Kline M, George R. In preparation. Defectors cannot be detected during 'small talk' with strangers.
72. Fowler KA, Lilienfeld SO, Patrick CJ. 2009 Detecting psychopathy from thin slices of behavior. *Psych. Assess.* **21**, 68–78. (doi:10.1037/a0014938)