ORIGINAL RESEARCH REPORT

Does Exposure to Hostile Environments Predict Enhanced Emotion Detection?

Willem E. Frankenhuis and Gijsbert Bijlstra

We used a Face-in-the-Crowd task to examine whether hostile environments predict enhanced detection of anger, and whether such enhanced cognition occurs for a different negative emotion, sadness, as well. We conducted a well-powered, preregistered study in 100 college students and 100 individuals from a community sample with greater exposure to hostile environments. At the group level, the community sample was less accurate at detecting both angry and sad faces than students; and, only students discriminated anger more accurately than sadness. At the individual level, having experienced more violence did not predict enhanced anger detection accuracy. In general, participants had a lower threshold (i.e., a more liberal criterion) for detecting emotion in response to anger than sadness. And, students had a higher threshold (i.e., a more conservative criterion) for detecting emotion than the community sample in response to both anger and sadness. Overall, these findings contradict our hypothesis that exposure to hostile environments predicts enhanced danger detection. Rather, our community sample was more prone to over-perceiving emotions, consistent with previous studies showing bias in threat-exposed populations. Future work is needed to tease apart the conditions in which people exposed to social danger show enhanced accuracy or bias in their perception of emotions.

Animals adapt to their environments on at least two timescales: across generations by natural selection, and within lifetimes through development. These processes are nested: natural selection shapes developmental mechanisms, which produce cognitive and emotional systems that tailor individuals to local conditions based on experience (Barrett, 2015; Belsky, Steinberg, & Draper, 1991; Chisholm, 1999; Ellis, Figueredo, Brumbach, & Schlomer, 2009; Frankenhuis, Panchanathan, & Barrett, 2013).

In this paper, we focus on hostile social environments, in which threats are common and severe. In such environments, mental systems might develop a focus on reducing the probability of harm (Belsky, 2008; Callaghan & Tottenham, 2016; Ellis & Del Giudice, 2014; Frankenhuis, Panchanathan, & Nettle, 2016; Taylor, 2006; Varnum & Kitayama, 2017). Reducing harm can be accomplished by enhancing cognitive abilities for detecting threats accurately (Frankenhuis & de Weerth, 2013), or by lowering the threshold for detection of threat, at the expense of increasing the number of 'false alarms' (Haselton et al., 2009; Haselton & Buss, 2000; Nettle & Haselton, 2006). Such 'erring on the side of

caution' might be adaptive when the costs of 'false alarms' are lower than the costs of 'misses,' depending also on the base rate of threats (Bateson, Brilot, & Nettle, 2011; McKay & Efferson, 2010). For instance, it might be costlier to infer that an angry person is calm than vice versa.

A cost asymmetry, however, might be absent or less distinct for sadness (inferring a sad person is happy), which entails lower risk of being harmed. Thus, even if the base rates of anger and sadness are similar, and higher in hostile environments, we expect thresholds for detecting anger to be generally lower than thresholds for detecting sadness. Variation in thresholds between negative emotions may be mechanistically possible, as previous research has shown differences in neural and cognitive processing of perceived anger and sadness (Blair, Morris, Frith, Perrett, & Dolan, 1999; Fox et al., 2000).

Previous studies of anger detection offer mixed results. Some studies show that growing up in hostile social conditions predicts enhanced accuracy in threat detection (reviewed in Ellis, Bianchi, Griskevicius, & Frankenhuis, 2017; Frankenhuis & de Weerth, 2013). Others report bias towards heightened sensitivity to threat (reviewed in Crick & Dodge, 1994; De Castro & van Dijk, 2017). For instance, physically abused children may orient more rapidly to angry faces and voices than controls do, and may be more accurate at identifying angry (but not other) facial expressions from degraded pictures (Pollak, 2008;

Behavioural Science Institute, Radboud University, Nijmegen, NL
Both authors contributed equally

Corresponding author: Willem E. Frankenhuis
(wfrankenhuis@gmail.com)

Pollak, Messner, Kistler, & Cohn, 2009). But these children may also exhibit a response bias toward anger, ascribing this emotion to situations where it is not fitting (Pollak, Cicchetti, Hornung, & Reed, 2000). Moreover, people who come from hostile social environments (e.g., violent families) are more likely to attribute hostile intent to ambiguous stimuli (e.g., neutral faces) than peers from safer environments (Crick & Dodge, 1994; De Castro & van Dijk, 2017).

In this study, we examine whether people who have experienced more violence are better at detecting threat, or overestimate threat, using a Face-in-the-Crowd task (i.e., FitC task; Hansen & Hansen, 1988); and whether this pattern is specific to anger or occurs for a different negative emotion, sadness, as well. We thus use sadness as a negative emotion to compare with anger. Whereas previous research has focused on well-delineated samples, such as children that experienced physical abuse (Pollak, 2008) or diagnosed with disruptive behavior disorders (De Castro & van Dijk, 2017), we assess a heterogeneous adult community sample. Our sample includes people who live in disadvantaged conditions for Dutch standards and who have experienced chronic (prolonged, intense) stress, such as exposure to violence, as well as individuals who are currently facing an acute stressor, but who have not experienced chronic stress. We compare this community sample with a lower-adversity sample, college students,[1] allowing us to assess accuracy and bias in emotion detection in the middle and lower range of adversity experiences, and to check the extent to which extant findings obtained with high-adversity samples generalize across different levels of adversity experience. We focus on the relationships between anger and sadness detection and (a) parental aggression, which in extreme cases may involve physical abuse, as well as (b) passive exposure to neighborhood violence, and (c) active involvement in violence.

At a group level, we expected the community sample to be more accurate at detecting an angry face in a crowd than the student sample (see Ellis et al., 2017; Frankenhuis & de Weerth, 2013). At an individual level, we expected people who had experienced more violence to be more accurate at detecting an angry face than people who had experienced less violence. We expected both of these relationships to be specific to anger ('anger superiority'; LoBue, 2009; Öhman, Lundqvist, & Esteves, 2001), rather than general to negative faces ('negativity bias'; Cacioppo, Gardner, & Berntson, 1999; Rozin & Royzman, 2001), because attention to danger tends to be more critical for survival and reproduction than detecting sadness.

Methods
We preregistered our sample size, materials, hypotheses, and statistical analyses at the Open Science Framework: https://osf.io/c8fne/. Our data is stored in the DANS repository and is accessible at: https://doi.org/10.17026/dans-zdh-ebhk (Frankenhuis, 2016). Our study was approved by the Ethics Committee, Faculty of Social Sciences, Radboud University; CSW2014-1310-250.
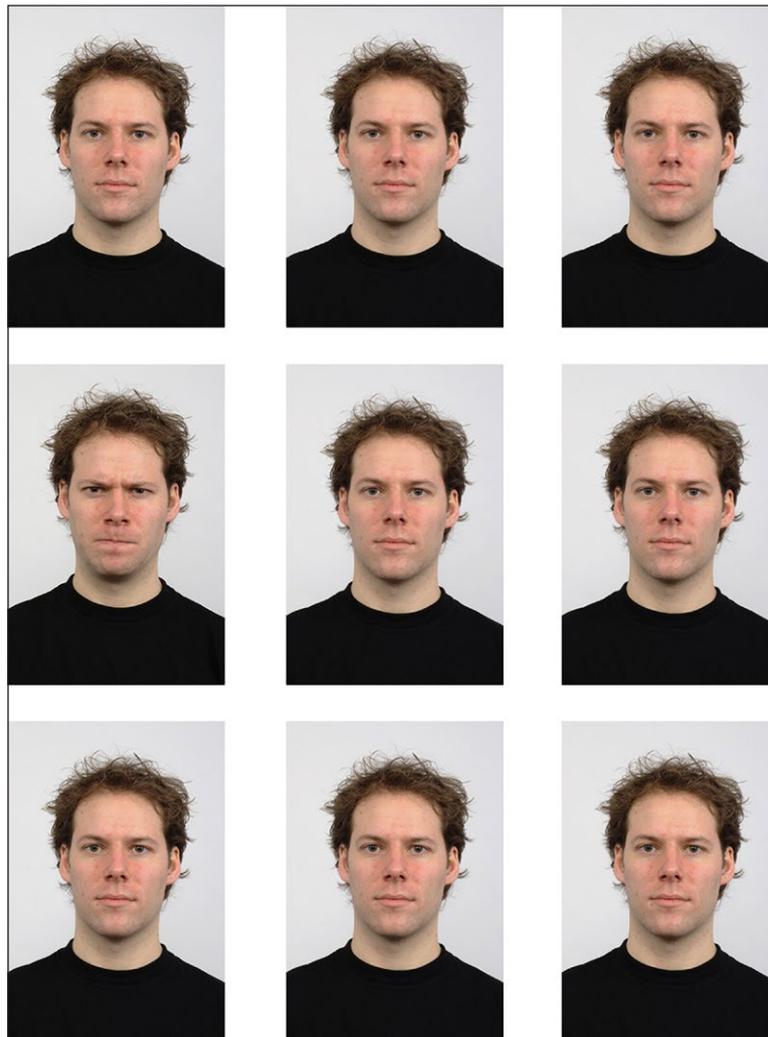
Participants
Our goal was to test 200 participants: 100 students and 100 from the community sample (see Preregistration). We initially tested 262 participants. After removing participants who had missing data—on the primary independent variables or the FitC task—we had 243 participants: 132 students and 111 from the community sample. We then excluded 11 participants from the community sample: two outliers (>3 SD on dependent variable), and nine for other reasons (e.g., poor vision, brain damage, drugged). Next, we excluded the last 32 students that we accidentally tested beyond our preregistered sample size, based solely on their date and time of participation, without having seen their data. The final sample thus comprised 200 participants:[2] 100 students (*M*age = 22.44, *SD* = 4.97, range: 18–61; 64 females) and 100 from the community sample (*M*age = 40.40, *SD* = 12.36, range: 18–65; 51 females).[3] Participants received € 10 or € 15 compensation, depending on whether their total session lasted 60 or 90 minutes.

Materials
Stimuli. We examined emotion detection using a novel FitC stimulus set, which we constructed using the Radboud Faces Database (Langner et al., 2010). We created 36 stimuli. A stimulus consisted of nine photos of one Caucasian male face, depicted in a 3-by-3 grid. In 18 stimuli, all faces showed a neutral expression (*neutral condition*). In nine stimuli, eight faces were neutral and one angry (*anger condition*, see **Figure 1**). In the other nine, eight faces were neutral and one sad (*sad condition*). Per emotion condition, the target emotion appeared once at each location in the 3-by-3 grid. Thus, 18 stimuli included one emotional expression (among eight neutral distractors). The other 18 showed only neutral expressions.

Neighborhood violence. We measured past (seven items; e.g., "In the neighborhood where I grew up, most people felt unsafe walking alone after dark") and current (seven items) exposure to neighborhood violence using the Neighborhood Violence Scale (for the development of this scale, see Frankenhuis, Roelofs, & de Vries, 2017). The subscales are identical except in referring to the past (<18 years) or present (current experiences). Participants rated items on a scale from 1–7 (*completely agree-disagree*). We computed a single score per participant by taking the mean over both subscales (α = 0.89).

Parental aggression. We measured parental aggression using two subscales (see Preregistration) of an abbreviated, Dutch version of the Parenting Questionnaire (see Ellis, Schlomer, Tilley, & Butler, 2012). These subscales, maternal aggression and paternal aggression, each consisted of four items. The items were statements (e.g., "My mother acted in a way that made me afraid that I might be physically hurt."). Participants rated the extent to which each statement described their childhood (0–16 years) on a scale from 1–5 (*never-always*). We computed a single parental aggression score per participant by taking the mean over both subscales[4] (α = 0.86). A higher score indicates greater paternal and maternal aggression. If

Figure 1: Face in the crowd example stimulus containing 8 neutral faces and 1 angry face (for privacy reasons, we depict a stimulus not used in the actual study).

participants had one parent, we used only that parent's subscale.

Involvement in violence. We measured active involvement in violence using a subset of four items (see Preregistration) from the Youth Risk Behaviour Survey (Eaton et al., 2012). Two items asked about the frequency of past involvement (14–17 years) and current involvement (in the last year) in a physical fight and needing treatment for injuries. Participants rated these on a scale from 1–5 (*0–6+ times*). The other two items just asked about the frequency of past and current involvement in a physical fight. Participants rated these on a scale from 1–8 (*0–12+ times*). These two response scales were identical to the ones used in the original Youth Risk Behaviour Survey, allowing us to compare samples across different studies. To be able to compute averages across scales with different numbers of response options (i.e., five or eight), we truncated the scores of 13 participants (6.5% of 200) who scored higher than five on the eight-option items, assigning them a value of five. We computed a single score per participant by taking the mean over all four items ($\alpha = 0.57$). This Cronbach's alpha is too low, implying the subset of results relating to this scale has low information value. With this caveat, we report our analyses as preregistered.

Procedure

We recruited the community sample via several organizations that help people who live in disadvantaged conditions for Dutch standards, facing such stressors as eviction and debt relief, unemployment, homelessness, previous incarceration, neighborhood and family violence.

The community sample completed the test-battery in a room at the respective community organization, and students in a test-cubicle at the university. All participants completed the same test battery, comprised of the present study and three other (non-related) studies, individually, in Dutch. Students were tested on a 24-inch desktop; the community sample on a 17-inch laptop.

Each trial started with a fixation-cross in the center of the screen (0.5s), followed by a stimulus (1.5s) and then a blank screen. Participants could respond only after the stimulus had disappeared, while viewing the blank screen. So, there was no tradeoff between accumulating evidence and responding sooner. By design, the response window was indefinite: the blank screen disappeared only when participants responded. We created our study to assess accuracy, not reaction times (see Preregistration). Nonetheless, we provide descriptive statistics about

reaction times in **Table 2**. We did not remove outliers based on reaction times.

During each trial, participants indicated whether an emotion was present or not by clicking the letter L or A, respectively. After a correct response, the next trial would start. After an incorrect response, the word "wrong" appeared (1s), before the next trial started. There was a four-trial practice-block before the test-block. The test-block consisted of the 36 grids, presented in random order, with each grid appearing only once.

Results

We use signal detection theory to analyze accuracy and bias (Green & Swets, 1974; Macmillan & Creelman, 2005). Participants may indicate an emotion when there is one (hit) or none (false alarm). We use the proportion of hits per emotion and a general false alarm rate (not emotion specific) to compute accuracy (d') and bias (c) using formulas provided by Stanislaw and Todorov (1999). Distinguishing between accuracy and bias is always useful, and absolutely necessary when the number of signal (emotion) and noise (neutral faces) trials is not equal. For instance, if 90% of all trials depict an emotion, participants with lower thresholds for detecting an emotion are more likely to respond correctly, by chance alone. In the extreme, a participant who estimates all trials to depict an emotion attains 90% correct, even when s/he is guessing at random. A measure of 'proportion correct' often confounds accuracy and bias (e.g., when the number of signal and noise trials is unequal due to missing data). Signal detection theory is a well-developed analytic method for describing decision-making in a wide variety of domains, including social perception (for an accessible introduction, see, Lynn & Barrett, 2014; Tan, Luan, & Katsikopoulos, 2017).

Parameter d' describes accuracy in discriminating signal (emotion) from noise (neutral faces), where higher d' implies greater accuracy, with lower bound zero. Criterion c describes the threshold for detecting the presence of a signal (i.e., an emotion): c equals zero implies no bias; negative c a lower threshold for detecting emotion (more liberal); and positive c a higher threshold (more conservative). Some participants attained extreme scores (e.g., 0% or 100% hits and/or false alarms). Hence, we used the log-linear method to improve estimates (Brown & White, 2005; Stanislaw & Todorov, 1999), applying this correction to all participants.

Preliminary analyses

In a different study of the same participant group, we have shown that our community sample self-reports having experienced higher levels of neighborhood violence and higher levels of harsh parenting, of which parental aggression is a subset, than our college students (see **Table 1** in Frankenhuis et al., 2017, for descriptive statistics, p-values, and Bayes Factors; at this link: https://osf.io/r2d67). Here, we show in addition that our community sample ($M = 1.48$, $SE = .06$) self-reports having been more involved in violence than our students ($M =$

1.15, $SE = .03$), $t(136.443) = -4.784$, $p < .001$, 95% CI [−.46, −.19], $r = .09$.[5]

We expected our three continuous predictors (i.e., parental aggression, passive exposure to neighborhood violence, and active involvement in violence) to be moderately or even highly correlated with each other. Therefore, we decided a priori to analyze them in separate models. Observed correlations ranged from .37 to .46, all $ps < .001$.

Primary analyses: Accuracy d'

A mixed between-within subjects ANOVA showed that the student sample ($M = 2.07$, $SE = .07$) was more accurate at detecting the presence of emotion than the community sample ($M = 0.88$, $SE = .07$), $F(1, 198) = 135.23$, $p < .001$, 95% CI [0.99, 1.40], $\eta^2 p = .41$; and participants were more accurate at detecting emotion in response to anger ($M = 1.57$, $SE = .06$) than sadness ($M = 1.39$, $SE = .05$), $F(1, 198) = 21.59$, $p < .001$, 95% CI [0.10, 0.25], $\eta^2 p = .10$.[6] However, these main effects were qualified by an interaction between emotion type and population, $F(1, 198) = 5.66$, $p = .018$, $\eta^2 p = .03$ (see **Table 1**).

Students were more accurate at detecting emotion in response to anger ($M = 2.21$, $SE = .08$) than sadness ($M = 1.94$, $SE = .08$), $F(1, 198) = 24.68$, $p < .001$, 95% CI [0.16, 0.37], $\eta^2 p = .11$. The community sample, however, was not more accurate at detecting emotion in response to anger ($M = .92$, $SE = .08$) than sadness ($M = .84$, $SE = .08$), $F(1, 198) = 2.57$, $p = .111$, 95% CI [−0.02, 0.19], $\eta^2 p = .01$. Comparing groups, students were more accurate than the community sample at detecting emotion in response to anger, $F(1, 198) = 135.38$, $p < .001$, 95% CI [1.07, 1.50], $\eta^2 p = .41$, and sadness, $F(1, 198) = 103.53$, $p < .001$, 95% CI [0.89, 1.32], $\eta^2 p = .34$.

Three separate between-within subjects ANOVAs, each adding the interaction between a standardized single continuous predictor and emotion type to the above model, revealed no interactions of emotion type with parental aggression, neighborhood violence, and involvement in violence, all $Fs < 1$, $p = $ ns.

Auxiliary analyses: Bias c

A mixed between-within subjects ANOVA showed that the community sample had a lower threshold for detecting the presence of emotion ($M = .12$, $SE = .03$) than students did ($M = .26$, $SE = .03$), $F(1, 198) = 14.39$, $p < .001$, 95% CI [0.07, 0.22], $\eta^2 p = .07$; and participants in general had a lower threshold for detecting emotion in response to anger ($M = .15$, $SE = .02$) than sadness ($M = .23$, $SE = .02$), $F(1, 198) = 21.59$, $p < .001$, 95% CI [−0.13, −0.05],

Table 1: Estimated marginal means and standard errors regarding accuracy (d') and bias (c) as a function of Emotion and Sample.

Sample	Accuracy d'		Bias c	
	Angry	**Sad**	**Angry**	**Sad**
Community	0.92 (.08)	.84 (.08)	.09 (.03)	.14 (.03)
Student	2.21 (.08)	1.94 (.08)	.20 (.03)	.33 (.03)

$\eta^2 p = .10$. However, these main effects were qualified by an interaction between emotion type and population, $F(1, 198) = 5.66$, $p = .018$, $\eta^2 p = .03$ (see **Table 1**).

Students had a higher threshold for detecting emotion in response to sadness ($M = .33$, $SE = .03$) than anger ($M = .20$, $SE = .03$), $F(1, 198) = 24.68$, $p < .001$, 95% CI [$-.19$, -0.08], $\eta^2 p = .11$. By contrast, the community sample showed no higher threshold for detecting emotion in response to sadness ($M = .14$, $SE = .03$) than anger ($M = .09$, $SE = .03$), $F(1, 198) = 2.57$, $p = .111$, 95% CI [-0.10, 0.01], $\eta^2 p = .01$.

Comparing groups, students had a higher threshold than the community sample for detecting emotion in response to anger, $F(1, 198) = 5.67$, $p = .018$, 95% CI [0.02, 0.19], $\eta^2 p = .03$, and sadness, $F(1, 198) = 19.60$, $p < .001$, 95% CI [0.11, 0.28], $\eta^2 p = .09$. Note that the within-subjects test statistics of this analyses are identical to those of the equivalent analysis of d', because the false alarm rate used in both analyses is not emotion specific.

Three mixed between-within subjects ANOVAs (as above) revealed no marginal predictive value of any of the continuous predictors, all $Fs < 1$, $p = ns$.

Discussion

Contrary to our predictions, students detected emotions more accurately than the community sample, even though our descriptive statistics show that both samples scored well above chance (**Table 2**). Consistent with anger superiority (LoBue, 2009; Öhman et al., 2001), participants more accurately perceived emotion in response to anger than sadness. This effect, however, was driven by the students. The task may have been more difficult for the community sample, which reduced this sample's accuracy on both emotions, and diminished the difference in its accuracy *between* emotions. Future work can evaluate this explanation by testing whether the interaction effect we observed replicates in an easier version of the task (e.g., when stimuli are displayed for longer).

Several mutually compatible factors might explain the community sample's lower accuracy. First, this sample erred on the side of caution by having a higher rate of 'false alarms' than students. As noted, such 'erring on the side of caution' can be adaptive when the costs of 'misses' are higher than those of 'false alarms' (Nettle & Haselton, 2006). For our community sample, not detecting anger or sadness, when it is in fact present, might be costlier than for our student sample; for instance, if these emotions are

Table 2: Means and standards error regarding response latencies in milliseconds and proportion correct as a function of Emotion and Sample.

Sample	Response latency		Proportion correct*	
	Angry	**Sad**	**Angry**	**Sad**
Community	1281 (94.76)	1344 (93.45)	.64 (.02)	.61 (.02)
Student	577 (34.64)	563 (32.41)	.83 (.01)	.74 (.02)

*One sample *t*-tests show that all proportions correct differ significantly from chance, all *p*'s < .001.

more likely to be followed by behaviors that could impose costs (e.g., requests for help, aggression). Alternatively, even if the costs of 'false alarms' and 'misses' would be identical for our samples, it is likely that the base rates of negative emotions is higher for our community sample, because the lives of people in harsher environments are more strained. It would be tremendously interesting for future work to empirically study base rates. This could be done either *objectively*, by quantifying the relevant statistics of people's natural environments (i.e., base rates and the extent to which emotional expressions predict particular behaviors), or *subjectively* by measuring people's expectations about such statistics (i.e., their priors about base rates and the predictive value of emotions). More generally, to better understand the development and utility of emotion perception styles, the field would benefit from greater investment in quantifying the actual statistics of people's lived developmental and current environments (e.g., Smith & Slone, 2017).

Second, the lab setting of students was more conducive to concentrating on the task than the on-site community setting. This difference was unavoidable: it was not feasible to bring the community sample into the lab. A more feasible possibility would be for a future study to bring an advantaged sample (e.g., students) to the community centers and test them there. It would be interesting to compare the performance of our samples when test settings are matched in this way. In such a study, the size of the computer screens should also be matched.

Third, students may have used a more effective scanning strategy. FitC tasks show multiple emotional expressions simultaneously in a single display. Participants may discriminate these expressions either by detecting emotions or based on lower-level features, such as the shape of eyebrows. Scanning based on lower-level features is known to improve performance in FitC tasks (Purcell & Stewart, 2010). It is possible, although to us not obviously plausible, that our students relied more on lower-level features than our community sample (note: this difference cannot explain the interaction effect we observed between emotion type and population).

Fourth, students might truly be better at detecting angry and sad faces than the community sample. If so, the question arises why. One possibility is that psychosocial adversity tends to impair emotion detection, except in people exposed to extreme violence (e.g., victims of physical abuse; Pollak, 2008), for whom detection of danger is a vital priority. However, our analyses at the individual level do not support this possibility, because in our study participants from more hostile environments did not display lower accuracy. Nonetheless, if students would have a general advantage in emotion detection, they should also be better at detecting positive emotions, such as happiness, than the community sample. Another possibility is that students are specifically better at detecting negative emotions (and equal or actually worse at detecting positive emotions) because they see fewer of these emotions in their daily lives (a novelty effect). We think this explanation is unlikely, however, because we suspect that all adults have had ample exposure to both

anger and sadness in real interpersonal interactions and via diverse forms of media (e.g., TV).

Our study's strengths include being well-powered, preregistered, and testing a large and heterogeneous sample, generating insight into a broad slice of the human experience. However, our study also has several limitations. First, as noted, the test settings inevitably differed between the student and the community sample. Second, testing our community sample in a computerized setting, rather than in a real-world, practical setting, may well have hindered their performance (Ellis et al., 2017). All of our students were familiar with test settings. Some of our community participants were not. Moreover, a few of them indicated feeling uncertain in a test setting (e.g., because they struggled in school). Third, as noted, we cannot rule out that some of our findings result from potential confounds such as group-level differences in scanning strategies, current stress levels, or motivation (note: if anything, our impression is that the community sample was highly motivated). Future research could explore the relationship between adverse experiences and accuracy in emotion detection using a more comprehensive set of adversity measures. Our observation that at the individual level hostile experience did not predict accuracy suggests that those experiences do not explain the community sample's lower accuracy.

Our study illustrates some of the challenges associated with conducting research with diverse populations in naturalistic settings, such as community centers, which afford less control than lab settings do. Emotion research is well known and appreciated for its effort to study cross-culturally and ethnically diverse populations (Barrett, Mesquita, & Gendron, 2011; Ekman, 1993; Elfenbein & Ambady, 2002). We hope that the current research of people with heterogeneous adversity experiences will connect with this tradition.

Data Accessibility Statements
Our data is stored in the DANS repository and is accessible at: https://doi.org/10.17026/dans-zdh-ebhk (Frankenhuis, 2016).

Notes
[1] We chose students as our low-adversity sample because we had better access to this population than to low-adversity non-student populations, and we anticipated, correctly, that assessing 100 high-adversity community participants would be a daunting, time-consuming task.

[2] Including all 232 participants in our preregistered analyses yields the same qualitative results.

[3] We had no hypotheses about age and gender, but report their analyses for completeness here. Including gender in our preregistered analyses yields the same qualitative results. Consistent with the extant literature (McClure, 2000), women detected emotions better than men. This sex difference was greater in the community sample than in the student sample, and was mainly due to men in the community sample performing less well. Including age in our preregistered analyses

also yields the same qualitative results. Accuracy levels decreased with age. With age included, the interaction between emotion type and population was marginally significant. This analysis is non-significant when all 232 participants are included. Also, we found no evidence for own-gender or own-race effects.

[4] Including only paternal aggression (i.e., without maternal aggression) in our preregistered analyses yields the same qualitative results.

[5] A one-sided Bayes Factor shows these data are 9038 times more likely to be generated by the hypothesis that our community sample has been more involved in violence than the data are to be generated by the null hypothesis (the two-sided Bayes Factor is 4519; JASP, 2016).

[6] Please note that we report the estimated marginal means and standard errors.

Competing Interests
The authors have no competing interests to declare.

Author Contributions
Both authors contributed equally.
- Contributed to conception and design: GB, WF
- Contributed to acquisition of data: WF
- Contributed to analysis and interpretation of data: GB, WF
- Drafted and/or revised the article: WF, GB
- Approved the submitted version for publication: GB, WF

References
Barrett, H. C. (2015). *The shape of thought: How mental adaptations evolve.* New York, NY: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199348305.001.0001

Barrett, L. F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*, 286–290. DOI: https://doi.org/10.1177/0963721411422522

Bateson, M., Brilot, B., & Nettle, D. (2011). Anxiety: An evolutionary approach. *The Canadian Journal of Psychiatry, 56*, 707–715. DOI: https://doi.org/10.1177/070674371105601202

Belsky, J. (2008). War, trauma and children's development: Observations from a modern

evolutionary perspective. *International Journal of Behavioral Development, 32,* 260–271. DOI: https://doi.org/10.1177/0165025408090969

Belsky, J., Steinberg, L., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: An evolutionary theory of socialization. *Child Development, 62,* 647–670. DOI: https://doi.org/10.2307/1131166

Blair, R. J. R., Morris, J. S., Frith, C. D., Perrett, D. I., & **Dolan, R. J.** (1999). Dissociable neural responses to facial expressions of sadness and anger. *Brain, 122,* 883–893. DOI: https://doi.org/10.1093/brain/122.5.883

Brown, G. S., & **White, K. G.** (2005). The optimal correction for estimating extreme discriminability. *Behavior Research Methods, 37,* 436–449. DOI: https://doi.org/10.3758/BF03192712

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology, 76,* 839–855. DOI: https://doi.org/10.1037/0022-3514.76.5.839

Callaghan, B. L., & **Tottenham, N.** (2016). The stress acceleration hypothesis: Effects of early-life adversity on emotion circuits and behavior. *Current Opinion in Behavioral Sciences, 7,* 76–81. DOI: https://doi.org/10.1016/j.cobeha.2015.11.018

Chisholm, J. S. (1999). *Death, hope and sex: Steps to an evolutionary ecology of mind and morality.* Cambridge, England: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511605932

Crick, N. R., & **Dodge, K. A.** (1994). A review and reformulation of social information processing mechanisms in children's adjustment. *Psychological Bulletin, 115,* 74–101. DOI: https://doi.org/10.1037/0033-2909.115.1.74

De Castro, B. O., & **van Dijk, A.** (2017). "It's gonna end up with a fight anyway": Social cognitive processes in children with disruptive behavior disorders. In: Lochman, J. E., & Matthys, W. (Eds.), *Wiley handbook of disruptive and impulse-control disorders,* 237–253. John Wiley & Sons Limited. DOI: https://doi.org/10.1002/9781119092254.ch15

Eaton, D. K., Kann, L., Kinchen, S., Shanklin, S., Flint, K. H., Hawkins, J., Lim, C., et al. (2012). Youth risk behavior surveillance – United States, 2011. *MMWR Surveillance Summary, 61,* 1–162.

Ekman, P. (1993). Facial expressions and emotion. *American Psychologist, 48,* 376–379. DOI: https://doi.org/10.1037/0003-066X.48.4.384

Elfenbein, H. A., & **Ambady, N.** (2002). On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychological Bulletin, 128,* 203–235. DOI: https://doi.org/10.1037/0033-2909.128.2.203

Ellis, B. J., Bianchi, J., Griskevicius, V., & **Frankenhuis, W. E.** (2017). Beyond risk and protective factors: An adaptation-based approach to resilience. *Perspectives on Psychological Science, 12,* 561–587. DOI: https://doi.org/10.1177/1745691617693054

Ellis, B. J., & **Del Giudice, M.** (2014). Beyond allostatic load: Rethinking the role of stress in regulating human development. *Development and Psychopathology, 26,* 1–20. DOI: https://doi.org/10.1017/S0954579413000849

Ellis, B. J., Figueredo, A. J., Brumbach, B. H., & **Schlomer, G. L.** (2009). Fundamental dimensions of environmental risk: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Human Nature, 20,* 204–268. DOI: https://doi.org/10.1007/s12110-009-9063-7

Ellis, B. J., Schlomer, G. L., Tilley, E. H., & **Butler, E. A.** (2012). Impact of fathers on risky sexual behavior in daughters: A genetically and environmentally controlled sibling study. *Development and Psychopathology, 24,* 317–332. DOI: https://doi.org/10.1017/S095457941100085X

Fox, E., Lester, V., Russo, R., Bowles, R. J., Pichler, A., & **Dutton, K.** (2000). Facial expressions of emotion: Are angry faces detected more efficiently? *Cognition & Emotion, 14,* 61–92. DOI: https://doi.org/10.1080/026999300378996

Frankenhuis, W. E. (2016). Does psychosocial adversity improve emotion detection? *DANS* [Data set]. Radboud University. DOI: https://doi.org/10.17026/dans-zdh-ebhk

Frankenhuis, W. E., & **de Weerth, C.** (2013). Does early-life exposure to stress shape or impair cognition? *Current Directions in Psychological Science, 22,* 407–412. DOI: https://doi.org/10.1177/0963721413484324

Frankenhuis, W. E., Panchanathan, K., & **Barrett, H. C.** (2013). Bridging developmental systems theory and evolutionary psychology using dynamic optimization. *Developmental Science, 16,* 584–598. DOI: https://doi.org/10.1111/desc.12053

Frankenhuis, W. E., Panchanathan, K., & **Nettle, D.** (2016). Cognition in harsh and unpredictable environments. *Current Opinion in Psychology, 7,* 76–80. DOI: https://doi.org/10.1016/j.copsyc.2015.08.011

Frankenhuis, W. E., Roelofs, M., & **de Vries, S. A.** (2017). Does exposure to social adversity enhance deception detection ability? *Evolutionary Behavioral Sciences.* Advance online publication. DOI: https://doi.org/10.1037/ebs0000103

Green, D. M., & **Swets, J. A.** (1974). *Signal detection theory and psychophysics* (reprint). New York: Krieger (original work published 1966).

Hansen, C. H., & **Hansen, R. D.** (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology, 54,* 917–924. DOI: https://doi.org/10.1037/0022-3514.54.6.917

Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuis, W. E., & **Moore, T.** (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition, 27,* 733–763. DOI: https://doi.org/10.1521/soco.2009.27.5.733

Haselton, M. G., & **Buss, D. M.** (2000). Error management theory: A new perspective in cross-sex mind reading.

Journal of Personality and Social Psychology, 78, 81–91. DOI: https://doi.org/10.1037/0022-3514.78.1.81

Haselton, M. G., & **Nettle, D.** (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and Social Psychology Review, 10*, 47–66. DOI: https://doi.org/10.1207/s15327957pspr1001_3

JASP Team. (2016). JASP (Version 0.8.4) [Computer software].

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & **van Knippenberg, A.** (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion, 24*, 1377–1388. DOI: https://doi.org/10.1080/02699930903485076

LoBue, V. (2009). More than just another face in the crowd: Superior detection of threatening facial expressions in children and adults. *Developmental Science, 12*, 305–313. DOI: https://doi.org/10.1111/j.1467-7687.2008.00767.x

Lynn, S. K., & **Barrett, L. F.** (2014). "Utilizing" signal detection theory. *Psychological Science, 25*, 1663–1673. DOI: https://doi.org/10.1177/0956797614541991

Macmillan, N. A., & **Creelman, C. D.** (2005). *Detection theory: A user's guide*. Mahwah, NJ: Lawrence Erlbaum.

McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin, 126*, 424–453. DOI: https://doi.org/10.1037/0033-2909.126.3.424

McKay, R. T., & **Efferson, C.** (2010). The subtleties of error management. *Evolution and Human Behavior, 31*, 309–319. DOI: https://doi.org/10.1016/j.evolhumbehav.2010.04.005

Öhman, A., Lundqvist, D., & **Esteves, F.** (2001). The face in the crowd revisited: A threat advantage with schematic stimuli. *Journal of Personality and Social Psychology, 80*, 381–396. DOI: https://doi.org/10.1037/0022-3514.80.3.381

Pollak, S. D. (2008). Mechanisms linking early experience and the emergence of emotions. *Current Directions in Psychological Science, 17*, 370–375. DOI: https://doi.org/10.1111/j.1467-8721.2008.00608.x

Pollak, S. D., Cicchetti, D., Hornung, K., & **Reed, A.** (2000). Recognizing emotion in faces: Developmental effects of child abuse and neglect. *Developmental Psychology, 36*, 679–688. DOI: https://doi.org/10.1037/0012-1649.36.5.679

Pollak, S. D., Messner, M., Kistler, D. J., & **Cohn, J. F.** (2009). Development of perceptual expertise in emotion recognition. *Cognition, 110*, 242–247. DOI: https://doi.org/10.1016/j.cognition.2008.10.010

Purcell, D. G., & **Stewart, A. L.** (2010). Still another confounded face in the crowd. *Attention, Perception, & Psychophysics, 72*, 2115–2127. DOI: https://doi.org/10.3758/BF03196688

Rozin, P., & **Royzman, E. B.** (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5*, 296–320. DOI: https://doi.org/10.1207/S15327957PSPR0504_2

Smith, L. B., & **Slone, L. K.** (2017). A developmental approach to machine learning? *Frontiers in Psychology, 8*, 2124. DOI: https://doi.org/10.3389/fpsyg.2017.02124

Stanislaw, H., & **Todorov, N.** (1999). Calculation of signal detection measures. *Behavior Research Methods, Instruments, & Computers, 31*, 137–149. DOI: https://doi.org/10.3758/BF03207704

Tan, J. H., Luan, S., & **Katsikopoulos, K. V.** (2017). A signal-detection approach to modeling forgiveness decisions. *Evolution and Human Behavior, 38*, 27–38. DOI: https://doi.org/10.1016/j.evolhumbehav.2016.06.004

Taylor, S. E. (2006). Tend and befriend: Biobehavioral bases of affiliation under stress. *Current Directions in Psychological Science, 15*, 273–277. DOI: https://doi.org/10.1111/j.1467-8721.2006.00451.x

Varnum, M. E., & **Kitayama, S.** (2017). The neuroscience of social class. *Current Opinion in Psychology, 18*, 147–151. DOI: https://doi.org/10.1016/j.copsyc.2017.07.032