Journal of Personality and Social Psychology
2011, Vol. 100, No. 4, 606–620

0022-3514/11/$12.00 DOI: 10.1037/a0022803

Community Perception: The Ability to Assess the Safety of Unfamiliar Neighborhoods and Respond Adaptively

Daniel Tumminelli O'Brien and David Sloan Wilson
Binghamton University

When entering an unfamiliar neighborhood, adaptive social decisions are dependent on an accurate assessment of the local safety. Studies of cities have shown that the maintenance of physical structures is correlated with the strength of ties between neighbors, which in turn is responsible for the crime level. Thus it should be theoretically possible to intuit neighborhood safety through the physical structures alone. Here we test whether people have this capacity for judging urban neighborhoods with 3 studies in which individuals observed photographs of unfamiliar neighborhoods in Binghamton, New York. Each study was facilitated by data collected during previous studies performed by the Binghamton Neighborhood Project studies. In the 1st study, observer ratings on neighborhood social quality agreed highly with reports by those living there. In the 2nd, a separate sample of participants played an economic game with adolescent residents from pictured neighborhoods. Players exhibited a lower level of trust toward adolescents from neighborhoods whose residents report lesser social quality. In the 3rd study, the maintenance of physical structures and the presence of businesses explained nearly all variation between neighborhoods in observer ratings (89%), whereas the specific features influencing play in Study 2 remained inconclusive. These and other results suggest that people use the general upkeep of physical structures when making wholesale judgments of neighborhoods, reflecting a adaptation for group living that has strong implications for the role of upkeep in urban environments.

Keywords: urban social behavior, prosociality, evolutionary psychology, disorder theory, environment perception

During social interaction, people condition their behavior on a variety of signals provided by the other individual, including facial features, posture, and cultural cues. In public places, interactions can be unpredictable, but the specific locale could provide additional information regarding those that might be expected. For example, the neighborhoods of a city vary in their level of safety, and it would be adaptive for an individual to use environmental cues to inform his or her social predisposition, being vigilant in a potentially dangerous neighborhood and relaxed otherwise. Prior knowledge would be helpful in formulating this response, but one entering an unfamiliar neighborhood would be dependent on the information provided by indirect signals.

Research on people's perceptions of novel environments—urban or otherwise—has identified two major factors that make scenes more appealing: (a) the ability to identify and understand the scene and (b) the curiosity and exploratory behavior it inspires (Kaplan, 1992). This work takes an evolutionary approach, describing such preferences as adaptations for an early human lifestyle, helping individuals to take appropriate paths when hunting and gathering. It does not consider, however, that group living may also have exerted an additional selection pressure on environment perception, leading it to evolve a socially oriented function similar to person perception. Just as an individual's appearance and mannerisms can signal aspects of his or her personality and quality as a social partner, the appearance of the streets, buildings, and open areas of a neighborhood reflects the treatment they receive from those who most often use them and indicate the quality of the social environment. When entering an area inhabited by others, individuals able to use such cues to make inferences about a community would be able to prepare appropriately for the types of interactions that can be expected there.

We refer to this proposed integration of environment and person perceptions as *community perception*, and in the studies presented in this article, we aim to demonstrate its existence as well as details of its function. We approach this topic with an evolutionary perspective, which is to say the focus is organized around two central ideas: (a) the environmental conditions that place selection

Daniel Tumminelli O'Brien, Department of Biology, Binghamton University; David Sloan Wilson, Department of Biology and Department of Anthropology, Binghamton University.

We would like to thank the National Science Foundation for funding this research through its support of Binghamton University's Evolutionary Studies Program and the Binghamton Neighborhood Project (Project No. 1046495). We would also like to recognize the assistance of the following individuals who were instrumental in the fruition of these studies: Tania O'Brien, for her photographic expertise; Jeffrey Carpenter, for advice on experimental economics; Andrew Gallup and Omar Eldakar, for assistance in organizing and running Studies 1 and 2; Matt Johnson, for helpful input during preparation; Mara Linek, Rada Maltseva, Alex Mand, Adam Scheer, Amanda Schneider, and Victoria Zdorodwski, for assistance in running Study 1; and Louis Alerte, Jeremy Cohen, Salvatore Ingrassia, Charles Norton, Sanjit Parhar, Thomas St. Pierre, and Nicholas White, for coding the physical features of individual photos.

Correspondence concerning this article should be addressed to Daniel Tumminelli O'Brien, Department of Biology, Binghamton University, Vestal Parkway East, PO Box 6000, Binghamton, NY 13902. E-mail: dobrien1@binghamton.edu

pressures on a trait and its function (ultimate mechanisms) and (b) the specific manner in which the resultant trait operates (proximate mechanisms). In a culture like our modern cities, distinct neighborhoods sit in proximity to one another, and a resident may travel through multiple communities in a day. If these communities vary in their safety level, there might be a selection pressure favoring those who have a capacity for community perception and use it to condition their social attitudes. To develop a more nuanced set of hypotheses regarding the trait's proximate mechanisms and how they interact with community variation, we turn to the extant literatures on person perception and urban criminology.

Personality Perception: Using the Available Cues

Humans have the tendency to quickly judge others on a variety of attributes. In keeping with the "thin slices" paradigm (Ambady, Bernieri, & Richeson, 2000), the assessment of many traits, including intelligence (Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004), sociosexuality (Gangestad, Simpson, DiGeronimo, & Biek, 1992), psychological disorders (Oltmanns, Friedman, Fiedler, & Turkheimer, 2004), and prosociality (Fetchenhauer, Groothuis, & Pradel, 2010; Verplaetse, Vanneste, & Braeckman, 2007), requires only a brief video lasting well under a minute. Depending on the personality trait, information as limited as a photograph of a person's face can be adequate to make an accurate judgment; what is necessary is that attributes correlated with the behavior in question be made available. For example, Gallup and Wilson (2009) found that body mass index was a reliable predictor of the level of intrasexual aggression perpetrated by high school girls. Similarly, independent raters appeared to use body mass index when estimating the aggressiveness of teenage girls in yearbook photos. Agreement between raters was considerable.

Even if a photo does not contain reliable indicators of the behavior in question, individuals still attempt to make such judgments. This has become particularly apparent in the literature on cheater (or, conversely, cooperator) detection, in which participants are asked to rate the trustworthiness or prosociality (i.e., tendency toward positive social behavior) of strangers. Individuals are able to discern cheaters from cooperators after an extended interaction (Brosig, 2002), or a short video (Fetchenhauer et al., 2010), but studies providing raters with only a photo find that predictions are no more accurate than chance. Only one study has violated this rule, and its protocol is distinct in that the photo was taken at the moment the pictured individual was choosing whether to cooperate or defect in an anonymous experimental economic game (Verplaetse et al., 2007). This suggests that neutral facial features do not signal prosociality but that the expressions produced in a social context can. Despite this tendency toward inaccuracy, interrater agreement is high in all such studies, implying that raters are relying on a specific set of facial characteristics that, though not correlated with prosociality, may be informative in some other way. Further research has demonstrated that these ratings are based on morphometrics that are associated with anger (Engell, Haxby, & Todorov, 2007; Oosterhof & Todorov, 2008). Although not accurately performing the task at hand, people have responded to cues that may be independently valuable in a social interaction.

The cheater detection literature illustrates how group living may have selected for the ability to quickly judge a stranger's quality as a social partner. Analogously, we propose that large-scale society—literally groups of groups—can promote the evolution of a trait that makes similar judgments about unfamiliar communities. Although this ability would have had to evolve very recently in human history, it seems feasible given the preexisting capacities for judging environments (environment perception) and social partners (personality perception). Community perception, then, would rely upon an algorithm that combines the relevant strengths of each, interpreting the environment primarily through cues that signal the local safety or trustworthiness. The cheater detection literature also provides an important lesson regarding the lack of such cues. If appropriate signals are not available, observers will be unable to accurately assess the safety of a neighborhood, despite the interest they may have in doing so. In this case, one may make such judgments inaccurately, relying on features that are informative in other ways, much as people interpret a seemingly angry facial expression as signaling a tendency to cheat.

Disorganization and Disorder: The Cause and Symptom of Unsafe Neighborhoods

Disorder theory (J. Q. Wilson & Kelling, 1982), also known as "broken windows" theory, posits that neighborhoods readily display their safety. Social disorder (e.g., public alcohol consumption) and physical disorder (e.g., overgrown vegetation) result when the local residents cannot or do not govern and maintain their community. These signals become crime attractors, as they indicate a safe haven for antisocial (i.e., delinquent) behavior. Residents tend to respond negatively to disorder, as surveys find that an individual's opinion of the neighborhood's social environment correlates highly with the amount of disorderly behavior the individual claims to observe there (Markowitz, Bellair, Liska, & Liu, 2001; Ross & Jang, 2000; Ross, Mirowsky, & Pribesh, 2001; Sampson & Raudenbush, 1999). Some have suggested that disorder itself might directly trigger fear, even if it is not accompanied by actually dangerous events, like assault (see Ross & Jang, 2000, for a review). During the 1980s, disorder theory became very popular among law enforcement, and New York City took a zero-tolerance approach to policing, hoping to discourage serious crime by stringently enforcing even the mildest of misdemeanors. There was a decrease in crime during the intervention, and one interpretation of this result (Corman & Mocan, 2005; Kelling & Sousa, 2001) claimed that the theory had been validated. An intriguing reanalysis found, however, that it was not the persecution of misdemeanors that best predicted the decrease in violent crimes but merely the police presence (Harcourt & Ludwig, 2006).

This instead provides support for a competing criminological theory, which states that social regulation of a neighborhood is necessary to prevent its infiltration by criminal elements (social disorganization theory; Shaw & McKay, 1969). Incidentally, this coincides with the evolutionary claim that cooperative systems must include a form of enforcement to prevent cheater strategies from succeeding (D. S. Wilson & Wilson, 2007). A rigorous test of the relationship between disorder, disorganization, and crime in Chicago found that disorder and crime are each symptoms of a neighborhood's inability to govern itself (i.e., collective efficacy; Sampson & Raudenbush, 1999), a finding that has since been

replicated in other locales (e.g., Kawachi, Kennedy, & Wilkinson, 1999). A neighborhood's safety, then, is a function of the social system constructed by its residents, a variable that must be assessed through surveys. As a proxy, the level of disorder, linked to the same deficiency as serious crime, would be a readily available indicator of local safety. This signal is present in physical structures in the form of poor maintenance, ranging from loose garbage to unkempt vegetation to damaged windows and doors.

The following studies test the dual hypotheses that individuals can accurately assess the safety of an unfamiliar neighborhood and do so by using indicators of disorder. In each, participants respond to photos of unfamiliar neighborhoods in Binghamton, New York. In the first study, participants report their opinions of a neighborhood's social interactions. We examine the accuracy of these inferences using ratings of the neighborhood's quality provided by residents. In the second study, an experimental economics game forces participants to make a social decision that has monetary impacts for both themselves and a partner living in the pictured neighborhood. This protocol includes a behavioral measure of trust, permitting us to measure its variation across contexts. In the third study, we attempt to determine the specific visible features that best explain the attitudes and behaviors witnessed in the first two studies.

Study 1

Testing the interrater consistency of judgments of different neighborhoods would be as simple as showing a set of participants a collection of images from a city. Testing the accuracy of these impressions, however, would require preexisting measures of social quality for those neighborhoods that participants are observing. This is available, as our study is part of a larger research program called the Binghamton Neighborhood Project (BNP), a collaboration between Binghamton University and community groups. The BNP uses an evolutionary perspective to unify multiple disciplines in the study of social behavior in urban contexts. In a previous BNP study, almost 2,000 six- to 12th-grade students at Binghamton High School responded to a survey that included questions about the relationships between one's neighbors (D. S. Wilson, O'Brien, & Sesma, 2009). These act as a measure of neighborhood social cohesion across the city, a reflection of social organization. Although the strength of a neighborhood's social organization correlates positively with income, we hypothesize that observers will assess a neighborhood's quality using signals of social disorganization (i.e., disorder), not indicators of income, as they are more proximately associated with safety.

In addition to providing measures of a community's social quality, the Developmental Assets Profile (DAP) contains a variety of questions about personal behavior, including drug use and self-esteem. As one might expect, many positive outcomes for youth correlate with a neighborhood's social quality, meaning that such characteristics might be accurately predicted with indicators of a neighborhood's social quality. There is an important distinction between individual tendencies and emergent properties of the community, however. A neighborhood's level of disorder is the result of publicly visible behaviors, like littering or failing to mow one's lawn, meaning it reflects the ability of the residential community to collectively enforce social norms. This says nothing of what the residents might do in private or in other social contexts. The physical structures are thus unlikely to include signals that are specific to these individual tendencies, leaving naive observers with no reliable information to facilitate predictions about them. In Study 1, we asked participants to estimate such characteristics of individuals, in addition to the social quality of the neighborhood. To independently test the accuracy of these two assessments, we intentionally used a subset of neighborhoods in which community quality does not correlate with individual tendencies. We hypothesize that observers will be unable to assess the behaviors of individual residents, and will base such judgments on disorder, conflating their impression of individuals with the quality of the community.

Method

Participants. Participants included 143 (45% male) undergraduate students who were enrolled in a biology–anthropology course at Binghamton University. Participants observed photos from neighborhoods across Binghamton and rated each on a collection of scales.

Materials. On September 23, 2007, digital pictures were taken at 20 semirandomly selected addresses from within the city of Binghamton. The initial list contained 25 randomly selected addresses and was pared down to 20 that maximized socioeconomic and geographic variation. At each of these addresses, four photos were taken: facing the address, looking across the street from the address, and looking each way down the street. When placed together, these approximated the visual experience of standing in the street in front of the address (see Figure 1 for examples). No photo included images of people.

Procedure. The collected images were shown to participants. For each of the 20 neighborhoods, the four pictures were displayed individually for 5 s each. Subsequently, all four images were displayed together (as seen in Figure 1) for 30 s. During this final 30 s, the participants were asked to rate the pictured neighborhood on the qualities listed below. Images from the next neighborhood were preceded by a slide noting its order number (e.g., Location 4). They were instructed to leave blank any address they believed they recognized. Fourteen addresses had at least one participant not respond ($M = 4.3$; minimum = 0, maximum = 32), making the final total 2,794 ratings.

Measures. While observing each neighborhood, participants rated two aspects of its social environment: the strength of ties between neighbors (*social cohesion*; two items, $\alpha = .84$) and their ability to govern the neighborhood (*social control*; three items, $\alpha = .87$). These items were on a 5-point Likert scale (1 = *strongly disagree*, 5 = *strongly agree*). For each respondent, a rating score for a neighborhood was calculated by summing the responses to all items and standardizing so that the lowest (all 1s) and highest (all 5s) possible scores were assigned values of 0 and 100, respectively. Owing to strong collinearity ($r = .88$, $p < .001$), we averaged scores on the two scales to form a measure of *social quality*. Field surveys in neighborhoods have found the correlation between these two measures to be of similar strength in vivo (Sampson, Raudenbush, & Earls, 1997).

In addition, three questions asked the participant to predict the attitudes of adolescents living in the neighborhood toward their own well-being, healthy habits, and prosociality. Consisting of only one item, these measures were left on the 5-point scale.

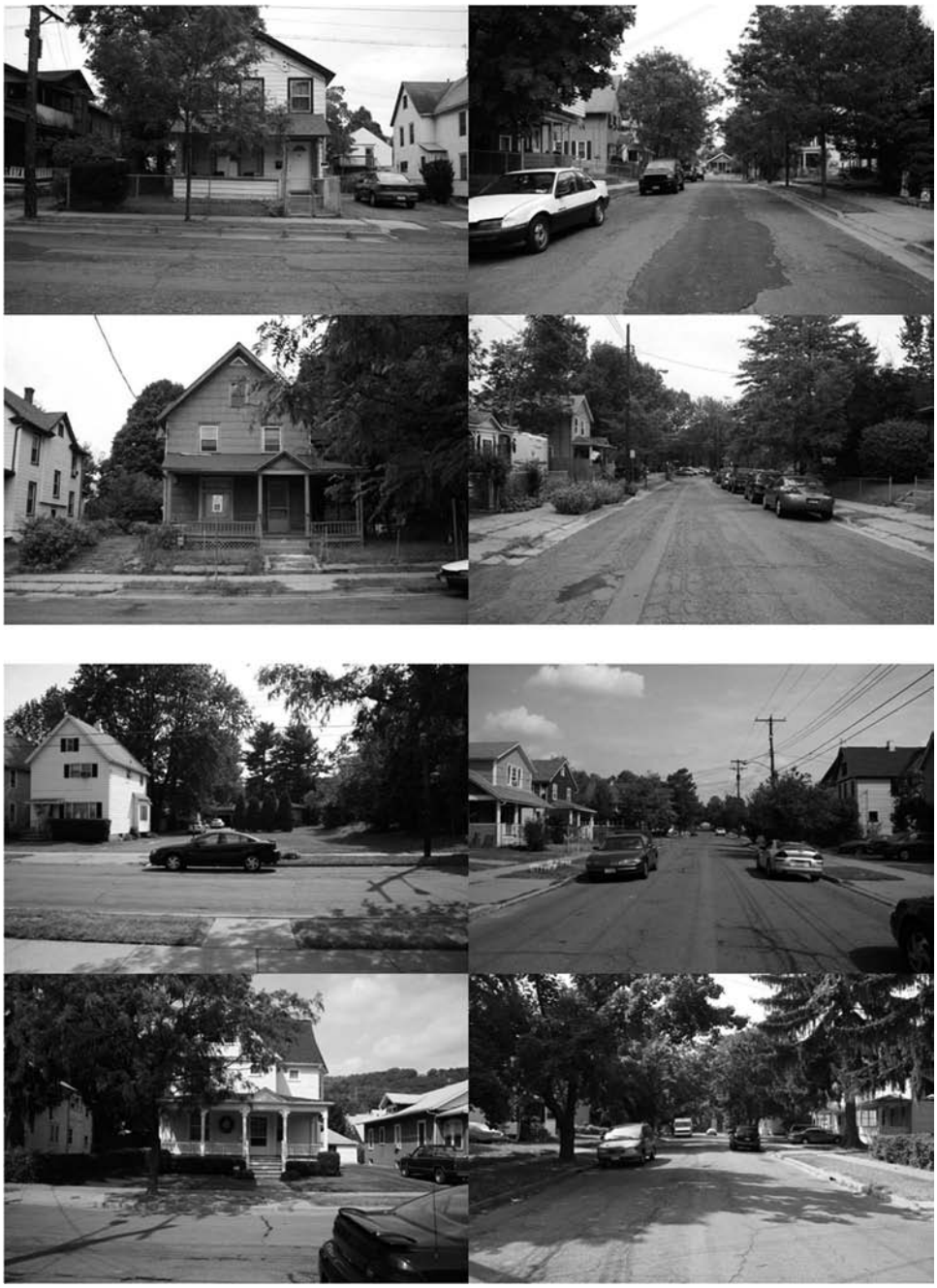

Figure 1. Example images from two neighborhoods as seen by participants.

Correlations between all rating categories at the response level are shown in Table 1; items are reported in the Appendix. Before the experiment began, participants reported how well they knew the geography of Binghamton on a 5-point Likert scale (1 = *not at all*, 5 = *extremely well*). This methodology was approved by Binghamton University's Human Subjects Resource Review Committee.

Neighborhood-level descriptors. In May 2006 nearly 2,000 Binghamton High School students in Grades 6–12 responded to the 58-item DAP, developed by Search Institute (http://www.search-institute.org) to assess the quality of life in adolescents (D. S. Wilson et al., 2009). Items from this were selected to form scales measuring well-being, healthy habits, and prosociality. Three DAP items reference the social quality of one's neighborhood. The items that raters responded to when viewing photos were drawn directly from these DAP scales. The additional questions regarding a neighborhood's social quality were crafted to closely resemble the rest of the scale both in word and spirit. The measures of well-being, healthy habits, prosociality, and social quality are on 0–100 scales as described above. To avoid confu-

Table 1
Correlations Between Scales Rated by Observers, at the Level of Individual Responses

Rating scale	1	2	3	4	5
1. Social cohesion	—	.88***	.80***	.71***	.75***
2. Social control		—	.79***	.72***	.73***
3. Well-being			—	.75***	.75***
4. Healthy habits				—	.75***
5. Prosociality					—

Note. $N = 2{,}791$.
*** $p < .001$.

Table 3
Correlations Between Neighborhood-Level Variables

Measure	1	2	3	4	5
1. Social quality	—	.36**	.58***	.65***	.76***
2. Well-being	.14	—	.51***	.52***	.29*
3. Healthy habits	−.12	−.59**	—	.81***	.45***
4. Prosociality	.05	−.27	.46*	—	.47***
5. Median income[a]	.54*	.29	−.11	−.03	—

Note. Whole-city correlations appear above the diagonal ($N = 63$), and those in the subset of photographed census block groups appear below ($N = 20$).
[a] Log-transformed.
* $p < .05$. ** $p < .01$. *** $p < .001$.

sion when discussing these scales, we denote in parentheses which measure (DAP or photo) is being referenced.

We used census block groups (CBGs) to approximate neighborhoods, in part because they permit the use of census statistics as independent variables. There are 63 CBGs in Binghamton, each intended to contain approximately 1,000 residents ($M = 752$, $SD = 228$). Using the mapping software ArcGIS (Version 9.6), we mapped responses to the DAP across the city and linked each student to his or her CBG of residence. We calculated neighborhood measures for each scale by averaging across the responses of all residents in a CBG. In addition, the median income of each CBG was accessed from the 2000 census. Owing to outliers, the variable was log-transformed before analyses. Descriptive statistics for each of these variables and correlations between them can be seen in Tables 2 and 3, respectively. These are reported for all 63 CBGs and the sample containing an address that was photographed.

Analysis. We used ArcGIS to link the 20 addresses and their photo ratings to the appropriate CBG (each was located in a different CBG), creating a design with responses nested within neighborhoods. To partition the variance associated with descriptors of raters (first level; e.g., one's knowledge of Binghamton) from descriptors of neighborhoods (second level; e.g., social quality [DAP]), we used hierarchical linear modeling (HLM 6.06; Raudenbush, Bryk, Cheong, & Congdon, 2004) to run multilevel regression models. We chose to use the parameters produced by unit-specific models, which focus on the variation across second-level units, rather than those that lean toward testing the population average of the entire sample. We also used traditional standard errors (as opposed to robust), as we did not expect responses to a neighborhood to be influenced by one another. HLM requires that there be no missing data, so three responses that left out an item were removed before analyses. The final sample size was 2,491 nested in 20 neighborhoods.

Results

Rating the social environment. An initial model, which was essentially a correlation between mean neighborhood ratings by residents and observers, found that adolescent residents and photo observers strongly agreed in their ratings of a neighborhood's social environment ($B = 1.35$, $d = 0.73$, $p < .001$; see Model 1 of Table 4 and Figure 2). In fact, the assessments of the two groups shared nearly 50% of their variation. In Model 2, we added other neighborhood descriptors as predictors (individual prosociality [DAP] and median income), and social quality (DAP) continued to be positively and significantly related to the ratings of photo observers ($B = 1.08$, $d = 0.61$, $p < .01$). Individual prosociality was included because of its theoretical relationship to community social quality, which is essentially a measure of reciprocal prosociality within the group. The magnitude of the parameter did decrease, but this is to be expected considering the collinearity between prosociality, median income, and social quality (DAP; see Table 3).

Models 3 and 4 first incorporated one's knowledge of Binghamton and then interactions between it and the neighborhood's social quality (DAP). In each a greater knowledge of Binghamton was associated with higher neighborhood ratings ($B = 1.28$, $d = 0.07$, $p < .001$, for latter model). Additionally, the significant positive parameter for the interaction effect between one's knowledge of Binghamton and the neighborhood's social quality (DAP) shows that those more acquainted with the city were better able to assess the social quality of a neighborhood, rating better neighborhoods higher and lower quality

Table 2
Descriptive Statistics for Neighborhood-Level Variables

	Whole city[a]			Subset[b]		
Measure	*M*	*SD*	Range	*M*	*SD*	Range
Social quality	54.66	11.23	11.11–80.56	53.29	8.07	36.56–66.41
Well-being	73.83	5.49	58.33–86.54	73.42	3.70	62.12–78.28
Healthy habits	70.49	7.09	41.67–88.89	70.89	4.77	63.89–81.82
Prosociality	61.12	6.51	33.00–77.93	61.59	4.45	54.87–70.00
Median income	$29,385	$16,465	$8,430–$90,143	$28,018	$11,980	$12,905–$59,567

[a] Includes all census block groups ($N = 63$). [b] Includes only the photographed census block groups ($N = 20$).

Table 4
Parameter Estimates From Multilevel Models Using Observer- and Neighborhood-Level Descriptors to Predict Ratings of Neighborhood Social Quality

	Model 1		Model 2		Model 3		Model 4	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*
Individual-level predictor								
Knowledge of Binghamton					1.31***	0.33	1.28***	0.03
Neighborhood-level predictors								
Social quality[a]	1.35***	0.32	1.08**	0.35	1.08**	0.35	1.08**	0.35
Median income[b]			18.87	15.56	19.81	15.54	19.76	15.56
Prosociality			0.93	0.54	0.93	0.54	0.93	0.54
Cross-level interaction								
Social Quality × Knowledge							0.09*	0.04
Approximate first-level R^2					.01		.01	
Approximate second-level R^2	.48		.54		.54		.54	

Note. $N = 2{,}791$ nested in 20 neighborhoods.
[a] Accessed from the 2006 application of the Developmental Assets Profile. [b] Log-transformed to maintain normality.
* $p < .05$. ** $p < .01$. *** $p < .001$.

neighborhoods lower ($B = 0.09$, $d = 0.04$, $p < .05$). Although each is statistically significant, their actual influence on ratings is limited. For example, according to Model 4, the maximum predicted difference in the rating of a neighborhood by someone reporting a 1 and another reporting a 5 for their knowledge of Binghamton is 4.7 points, which is only half a standard deviation in the subset of neighborhoods used here (see Table 2).

Rating other features of the neighborhood. Following the above results, we ran three models, one predicting each of the other three photo ratings: well-being (photo), healthy habits (photo), and prosociality (photo). Each model included six variables: the corresponding measure from the DAP, one's knowledge of Binghamton, and an interaction between the two; the measure of neighborhood social quality provided by the DAP and an interac-

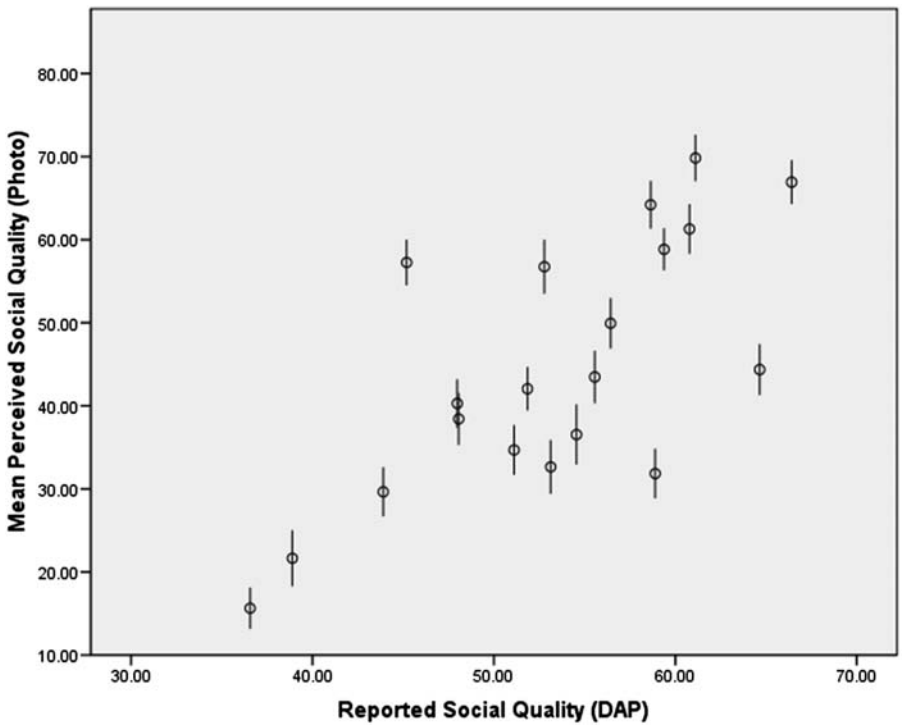

Figure 2. Relationship between social quality as reported by resident adolescents and as rated by participants viewing photos of the neighborhood. Error bars reflect 95% confidence interval for the mean of each neighborhood's photo-based ratings. DAP = Developmental Assets Profile.

tion between that and one's knowledge of Binghamton; and median income. The findings can be almost completely generalized across the three measures (see Table 5):

1. People did not accurately assess the level of well-being, healthy habits, or prosociality in the local youth, as indicated by the lack of association between their ratings and the measurements taken from the DAP.

2. These ratings appear to be based almost completely on those cues being used to assess the social environment. Social quality (DAP) was the most accurate predictor of the average response to a neighborhood's photos for all three measures. The magnitudes of these relationships were nearly identical to the one with social quality as rated from photos.

3. Observers did not begin using signals of income to inform their ratings when attempting to assess these three qualities, but relied throughout on indicators of the social environment.

4. Finally, the results regarding the effect of one's knowledge of Binghamton on ratings were mixed. In general, it seemed that those more acquainted with Binghamton viewed the neighborhoods more favorably, although this was not the case for ratings of prosociality. Also, it seemed that these people may assess more accurately, following some set of cues in the photos, although, as above, the effect sizes are small and may be an artifact of a large sample size.

Discussion

Viewing only pictures of a neighborhood's physical structures, the participants were able to accurately judge the social dynamics of a neighborhood. Nearly 50% of this variation coincided with the ratings provided by neighborhood residents, showing that these predictions are not just better than chance but quite consistent with reality. Further, raters were not responding to cues of affluence but other, unidentified features in the photos. It seems observers used these same features when asked to judge the lifestyles of individual residents. Even ratings of individual prosociality were a reaction to indicators of the neighborhood's social quality. By decoupling the correlations between individual behaviors and community characteristics, we found that the pictures were not providing information that is specific to the behaviors of individuals—at least not that the observers were accessing. Instead, participants appeared to be using a set of characteristics to generate all judgments regarding the neighborhood, even though these characteristics are reliable only as indicators of the community's social dynamics. Despite being potentially inaccurate, this unidimensional assessment of both community quality and individual behavior seems both efficient and adaptive when entering an unfamiliar neighborhood and preparing to interact with the environment.

It is important to note that the lack of correlation between the features of a community and individual behaviors was part of our experimental design and not necessarily the case in actual cities (see, e.g., Table 4). This simply serves to demonstrate that the signals visible in a community's physical structures are only informative about collective processes and do not provide direct indicators of individual tendencies. In the case that the attributes of the community and its individual residents were correlated, individuals might make accurate assumptions about individual behavior but only by virtue of this association.

There is some suggestion that the ability to interpret the physical structures has a cultural component, as a greater knowledge of Binghamton led to higher and more accurate ratings. The effect sizes behind these associations were small, however, and even individuals with little to no experience with the city were able to make accurate judgments. The limited cultural variation in the participants—Binghamton University is a public university with students hailing primarily from New York State—may have masked a more extensive cultural variation in community perception.

Table 5
Parameter Estimates From Multilevel Models Testing the Effect of Observer- and Neighborhood-Level Descriptors on Ratings of Neighborhood Social Quality

	Well-being		Healthy habits		Prosociality	
Variable	*B*	*SE*	*B*	*SE*	*B*	*SE*
Individual-level predictor						
Knowledge of Binghamton	0.03*	0.015	0.03*	0.016	0.002	0.016
Neighborhood-level predictors						
Well-being[a]	−0.04	0.028				
Healthy habits[a]			0.01	0.02		
Prosociality[a]					0.03	0.02
Social quality[a]	0.05**	0.015	0.04*	0.015	0.04***	0.01
Median income[b]	0.92	0.68	0.72	0.65	0.72	0.65
Cross-level interactions						
Well-Being × Knowledge	0.01**	0.004				
Healthy Habits × Knowledge			−0.005	0.004		
Prosociality × Knowledge					−0.01**	0.004
Social Quality × Knowledge	0.01*	0.002	0.003	0.002	0.008***	0.002
Approximate first-level R^2	.00		.00		.01	
Approximate second-level R^2	.50		.44		.55	

Note. $N = 2{,}791$ nested in 20 neighborhoods.
[a] Accessed from the 2006 application of the Developmental Assets Profile. [b] Log-transformed to maintain normality.
* $p < .05$. ** $p < .01$. *** $p < .001$.

Study 2

When interacting with others, the effectiveness of a social strategy is in great part dependent on the trustworthiness of one's partner. Fittingly, people are overwhelmingly more likely to cooperate with strangers they believe to be prosocial (e.g., Brosig, 2002). It has often been proposed that the evolved function of attitudes and emotions is to promote specific action tendencies that produce behaviors adaptive for the immediate circumstance (Frijda, 1989; Lazarus, 1991; Levenson, 1994; Tooby & Cosmides, 1990). In this case, the ability to judge the safety of a neighborhood from its physical structures is relevant only insofar as it influences one's approach to social interaction. Here we used a game developed by experimental economists called the Sequential Prisoner's Dilemma (SPD) to test the hypothesis that people use the information embedded in the physical structures of a neighborhood to elect appropriate behaviors. In this game, a participant must make decisions that will impact monetary payoffs for him- or herself and a social partner. Such protocols are especially effective in that they simplify interactions, creating easily interpreted measurements of social behavior. Further, the real-world implications associated with monetary payoffs qualify them as performance-based measures of social behavior that are less vulnerable to the effects of social desirability than standard surveys. If participants condition their behavior on the physical appearance of a neighborhood, their choices when asked to play with a local resident will demonstrate how they would behave in such an environment. Again, results from previous BNP studies allowed us to assess the ability of participants to respond appropriately.

Method

Participants. Participants included 34 (56% male) undergraduate and graduate students in a biology–anthropology course at Binghamton University. All participants observed photos from neighborhoods across Binghamton, playing the SPD as if with a resident from each. No participant included in Study 1 was a part of Study 2.

Materials. In the two-player SPD, each player can either cooperate or defect, with two cooperators each receiving a greater amount than two defectors (in this version, \$30 vs. \$15) but a mixed interaction resulting in a lower payoff for the cooperator than the defector (\$10 vs. \$45). A first mover chooses whether to cooperate or defect, enabling the second player to choose on the basis of the first player's decision. A player may have to choose between cooperation and defection in one of three circumstances: as a first mover, a second mover with a cooperator, or a second mover with a defector. Electing to cooperate as a first mover ("offers of cooperation") is an indicator of trust, as the second mover has the opportunity to exploit cooperation. As a second mover, cooperating with a cooperative first player is a measure of "reciprocation," and cooperating with a defector is an indicator of "self-sacrifice."

Procedure. Participants were shown the images from nine of the 20 neighborhoods used in Study 1. Initially, we generated a random selection of nine neighborhoods, then slightly modified it to expand variation in social quality across neighborhoods. Again, for each neighborhood, the participants saw each of the four images alone for 5 s each and then all four together (as seen in Figure 1) for 30 s. During this longer period, participants were asked to play the SPD as if they were playing with an adolescent living in the neighborhood. They were assured that this would occur and that real monetary payoffs would be given (see below). Roles (i.e., first or second player) were to be assigned randomly, so they must provide responses for each of the three possible situations. Participants were asked not to play with a given neighborhood if they believed they recognized it. For each neighborhood, at least one participant left responses blank ($M = 1.3$; minimum = 1, maximum = 2). The remaining sample size was 294 responses nested in nine neighborhoods.

The participants were notified that only two individuals, chosen at random, would actually play the game for money. One would act as a first mover, the other as a second mover, and each would play with a resident of one of the photographed neighborhoods.[1] This methodology was approved by Binghamton University's Human Subjects Resource Review Committee.

Measures. For each neighborhood, a participant indicated whether or not he or she would offer cooperation, reciprocate cooperation, and self-sacrifice. These were coded as dichotomous variables (1 = cooperate, 0 = defect). The first two measures allowed us to analyze how trust and reciprocation vary with the appearance of a social partner's neighborhood. We did not analyze the third measure, as it is very rare that a second mover chooses to cooperate with a noncooperator, essentially sacrificing \$5 and giving the other individual \$30 more. Participants also rated their knowledge of Binghamton as in Study 1.

Analysis. Again, a nested design was created, with responses to a set of photos linked to the appropriate address and, in turn, the associated CBG. HLM was used to test logit models, owing to the dichotomous nature of the two dependent variables, offering cooperation as a first mover and reciprocating cooperation as a second mover. In this study we used unit-specific models but with robust standard errors, owing to a severe disagreement between the results given by traditional and robust standard errors. In this case, it is most appropriate to use the robust standard errors because the distribution of random errors across second-level units is not normally distributed. We hypothesize that this is because there are obvious cues that can inform behavior in the photos of neighborhoods with very high and very low social quality, but those neighborhoods rated centrally are less likely to elicit consistent decisions when the outcome variable is dichotomous, as are behaviors in the SPD. This would skew random effects to correlate with an independent variable, violating the assumptions of the model. To check the robustness of these findings, we also ran linear regressions using neighborhood descriptors to predict the proportion of individuals who offered cooperation or reciprocity when playing with each neighborhood.

[1] This was accomplished with data from a previous study at Binghamton High School in which students played the SPD (O'Brien et al., 2007). For the nine addresses whose images were used in Study 2, a high school participant who lived within the CBG was selected. The students' behavior in the game was then used to simulate the social interaction described to the Binghamton University students. Two of the neighborhoods were chosen at random, one for each randomly selected Binghamton University student. The responses of the high school representative filled the necessary role (first or second mover). Money was distributed to the university participants as per the rules of the game.

Results

The relative success of offering cooperation as a first-mover is influenced by the likelihood that it will be met with reciprocation; thus we assume that variation in play is based on the extent to which participants trusted residents of each neighborhood. For every point that a neighborhood's social quality (DAP) increased, a participant was 5% more likely to offer cooperation, meaning that the images of these neighborhoods elicited greater trust ($B = 0.05$, odds ratio [*OR*] = 1.05, $p < .01$; see Model 1 of Table 6), regardless of the university participant's own age, sex, or knowledge of Binghamton. Median income was added in Model 2, and increases in a neighborhood's social quality (DAP) remained predictive of trusting behavior, though the effect size was diminished. This was not surprising owing to the strong correlation between the two variables in this subset of neighborhoods ($r = .61$, $p < .10$). Median income was a nonsignificant predictor. Because the assumptions of HLM were violated, we checked the accuracy of the analysis by replicating this last model in the form of a standard multiple regression. This model corroborated the results from HLM, with a neighborhood's social quality (DAP) positively predicting the proportion of individuals exhibiting trust when playing with a particular neighborhood ($B = 0.73$, $p < .05$). Again, median income was a nonsignificant predictor.

Second, we analyzed the tendency of individuals to reciprocate when playing with a cooperator, also reported in Table 6. The first model (Model 3) found that people were more likely to reciprocate when playing with a resident from a pictured neighborhood with higher social quality (DAP). The inclusion of median income in Model 4, however, caused the strength of the parameter to shrink to a level of marginal significance. This was also the case for the standard regression ($B = 0.54$, $p < .10$), although a stepwise version of the regression found a neighborhood's social quality (DAP) to be the main predictor of the proportion of people who reciprocated when playing with an adolescent from a particular neighborhood ($B = 0.72$, $p < .05$; cut-point, $\alpha = .05$). Interestingly, the most consistent predictor of reciprocation was one's knowledge of Binghamton ($B = 0.33$, $OR = 1.38$, $p < .01$), an increase of 1 point on this scale being associated with a 38% greater chance of reciprocating. There was no effect of a participant's age or sex on behavior.

Variation in first-mover behavior (i.e., trust) was more closely associated with differences in the photos than responses to cooperators. This speaks to the different considerations that go into each decision, as the first involves an assessment of a partner's trustworthiness. As in Study 1, it appears that participants are using information in the photos to make such judgments. Although this assessment would not be necessary when determining whether to reciprocate, we still see a relationship between cooperation and variation in the pictured neighborhoods.

Discussion

Considerable research has found that setting has a strong influence on social attitudes, as individuals are more likely to exhibit prosocial behavior when in rural environments than urban ones (see Steblay, 1987). An earlier BNP study (D. S. Wilson et al., 2009), however, found similar variation to exist across urban environments using the "lost-letter" method (Milgram, Mann, & Harter, 1965). Stamped envelopes, addressed to a specific location, were dropped at randomly selected locations throughout the city of Binghamton. Envelopes arriving through the mail represented a small act of prosociality, as they had been picked up by passersby and put it in a mailbox. Again, setting was more responsible for prosociality than individuals, as return rate was positively associated with a neighborhood's level of social cohesion, rather than the generalized level of prosociality reported by individual residents.

Similarly, the participants in this study acted in a distrusting manner if neighborhoods appeared unwelcoming, reacting adaptively to signals of the quality of the social environment. Again, they did not respond to cues of median income, despite the correlation between the two variables. Interestingly, in the experimental economics study of Binghamton adolescents that was used to facilitate the experiment (see footnote), the quality of the local social environment did not directly predict trustworthiness of residents, although population density, which tends to diminish the ability of a community to govern itself, was associated with a lower rate of reciprocation (O'Brien, Wilson, Eldakar, & Carpen-

Table 6
Parameter Estimates From Multilevel Logit Models Using Observer- and Neighborhood-Level Descriptors to Predict the Likelihood of Cooperating When Playing the Sequential Prisoner's Dilemma With a Resident of a Pictured Neighborhood

	Offers of cooperation (trust)						Reciprocation					
	Model 1			Model 2			Model 1			Model 2		
Variable	*B*	*SE*	*OR*	*B*	*SE*	*OR*	*B*	*SE*	*OR*	*B*	*SE*	*OR*
Individual-level predictors												
Knowledge of Binghamton	−0.04	0.10	0.96	−0.04	0.10	0.96	0.32**	0.10	1.38	0.33**	0.10	1.38
Age	0.15	0.09	1.16	0.15	0.09	1.16	0.02	0.10	1.02	0.02	0.10	1.02
Sex[a]	0.03	0.20	1.03	0.03	0.20	1.03	0.08	0.24	1.09	0.09	0.24	1.09
Neighborhood-level predictors												
Social quality[b]	0.05**	0.02	1.05	0.03**	0.01	1.03	0.03***	0.007	1.03	0.02†	0.01	1.02
Median income[c]				0.90	0.54	2.46				0.58	0.40	1.78

Note. $N = 294$ nested in nine neighborhoods. *OR* = odds ratio.
[a] Female = 1. [b] Accessed from the 2006 application of the Developmental Assets Profile. [c] Log-transformed to maintain normality.
† $p < .10$. ** $p < .01$. *** $p < .001$.

ter, 2007). As seen in Study 1 and mentioned above, the cues that informed participant reactions to a neighborhood were not specific to the behaviors of individual residents but to the quality of the environment they had constructed and the dangers it may harbor. From this assessment, one derives a level of trust that will be most effective in navigating the general setting, which is then projected upon all individuals therein.

In the case of responding to a cooperator (i.e., reciprocation), logic states that there is no adaptive purpose in varying behavior across neighborhoods, as the social partner has already acted. However, participants reciprocated more when interacting with certain neighborhoods than others, and the primary determinant of this variation again seems to be the local social quality, although the results are less conclusive in this case. This further supports the theory proposed in Study 1 that participants observing the images experience a generalized emotional response to the pictured neighborhood, judging it in a wholesale manner and reacting accordingly. One caveat to this interpretation, however, is that participants may be fabricating variation in response to the experimental setting itself.

If reliable, these results offer support for the theory of strong reciprocity, which states that humans have evolved a general tendency for prosocial behavior through generations of group-based survival (Bowles & Gintis, 2004; Gintis, 2000). This predisposition for cooperation is attuned to the mutual affinity and norms one shares with a social partner, regardless of whether future interactions are expected. In this case, we see reciprocation based on the same criteria as offering cooperation; once one determines that a neighborhood's residents are worthwhile social partners, one sees them as both trustworthy and worth cooperating with, forgoing the temptation to defect. On the other hand, when observing a neighborhood that appears to harbor untrustworthy individuals, players often decided not to invest socially in the residents by reciprocating. Further evidence for strong reciprocity comes from the relationship between one's knowledge of Binghamton and the tendency to reciprocate. Although it is generally accepted that individuals tend to cooperate more often with family members or close friends, here people are seen as becoming more likely to exhibit prosociality toward complete strangers on the basis of nothing more than shared familiarity with a city and, presumably, its culture.

Study 3

Studies 1 and 2 have shown that individuals are able to intuit the quality of a neighborhood's social environment by viewing its physical landscape and that they react adaptively but give no indication of how they are doing this. As mentioned above, the bonds between neighbors are an instrumental resource for the informal governance of a neighborhood, and where there is limited social organization crime is more likely to occur (Sampson et al., 1997). Ironically, a direct measurement of these social relationships would require interaction with residents—examples of the exact experiences for which a passerby would need such information. There is a strong correlation, however, between neighborhood governance, crime, and physical indicators of disorder (Markowitz et al., 2001; Sampson & Raudenbush, 1999). Some claim that these signals directly attract crime, as would-be criminals recognize an area where delinquent behavior entails little risk of punishment (Corman & Mocan, 2005; Kelling & Sousa, 2001; J. Q. Wilson & Kelling, 1982).

Residents who perceive their neighborhood as being more disorderly also report their neighborhood as being more dangerous, regardless of whether they have witnessed or been a victim of a serious crime (Markowitz et al., 2001; Ross & Jang, 2000; Ross et al., 2001). If physical evidence of poor maintenance influences the attitudes and behaviors of those living there, it would seem logical that those with less a priori knowledge would base their own attitudes and behaviors on the same cues. In this study we attempted to establish which forms of disorder—be they deterioration of the houses, the pavement, the lawns, or other elements of a neighborhood—and other visible items, like lawn and house decorations, are the best indicators of the social relationships shared by residents. In turn, we tested the hypothesis that the capacity for community perception seen in Studies 1 and 2 is a response to visible disorder and attempted to identify those specific features in an image that are most responsible for these judgments.

Method

Measures. Seven individuals objectively rated each image on the physical features noted in Table 7. They rated 100 images—the 80 that were used in Study 1 and 20 dummy images—in a set order. The dummy images comprised the first 20 to allow individuals to solidify their rating system before rating those images relevant to the study. The other 80 images were randomized in such a way that raters did not know which images came from the same neighborhood. The raters knew that all images were randomized but not about the 20 dummy images.

The items rated were meant to reflect the care invested in a neighborhood's physical structures (i.e., the level of physical disorder). Ratings were left blank if nonapplicable for a given image (e.g., rating lawn quality in a picture of a street). Ratings were consistent for those variables that permitted some level of subjectivity (e.g., "Are the exteriors well painted?"; see Table 7), permitting the averaging of all ratings to create an image-specific score. These scores were then averaged across the four images taken at each address to create an address-specific score. As for

Table 7
Physical Characteristics of Neighborhood Images Assessed by Independent Raters and Interrater Reliabilities

Item	α
Are the exteriors well painted?	.91
Are the driveways well cared for?	.81
Is the grass appropriately trimmed/mowed?	.87
Is other vegetation appropriately trimmed?	.90
Are there any "junk" vehicles in the street/driveway?[a]	—
Are the lawns/streets kept clean of garbage?	.83
Are there any buildings that look abandoned?[a]	—
Are there any broken windows or doors?[b]	—
Is the sidewalk in good repair?	.78
Are the streets cracked or unevenly paved?	.96
Is there any graffiti present?[b]	—
Are there any lawn/porch decorations?	.81
Are there any businesses?[a]	—

Note. $N = 80$ images rated by seven individuals. All were rated on a 5-point scale unless otherwise noted. Dashes indicate that Cronbach's alphas were not calculated for variables with no variance across raters.
[a] 1 = yes, 0 = no. [b] Three-point scale (*none, some, much*).

more objective variables, no images contained broken windows, "junk" vehicles, or graffiti, and only one contained an abandoned building, leading us to discard these variables before analysis. Fifteen images from seven neighborhoods contained businesses, providing enough variation to include it as a dichotomous variable in analyses (*business*; 1 = business visible from address).

Analysis. Those variables rated on a 5-point scale had considerable shared variation (see Table 8). To reduce these to simpler factors, we ran a principal component analysis (PCA). We elected this form of extraction over a factor analysis because it analyzes all variation, not just overlapping variance. This seemed appropriate, as the variables are naturalistic observations. We assumed correlations between any resultant factors and chose a varimax rotation. The PCA was run at the image level, not the neighborhood level. This is because 20 neighborhoods would be an inadequate sample size to run a PCA. Instead, each image included in Study 1 (80 in total, four from each neighborhood) was treated as a single case. Although the sample size is still small, this can be acceptable when component loadings are high, as they are here (Guadagnoli & Velicer, 1988; Sapnas & Zeller, 2002). The results of the PCA informed the creation of variables that, along with the presence of a business, acted as neighborhood descriptors in multilevel models predicting participant responses to neighborhoods in Studies 1 and 2. Each model was run in HLM and used the same parameters and standard errors as the corresponding model in the previous studies (see above).

Results

Quantifying disorder. As mentioned above, the PCA was run at the image level, including all images used in Study 1 ($N = 80$). For images that did not contain lawns or driveways, we imputed ratings by using the average rating for the neighborhood's other images. The two resulting components reflect the amount of care invested in house (ratings of paint, grass, vegetation, garbage, and decorations) and pavement (ratings of driveway, sidewalk, and streets; see Table 9 for complete results). Notably, the appearance of the grass had the greatest shared variance with the quality of other physical features in an image and was the same feature that most highly correlated with a neighborhood's social environment ($r = .73$, $p < .001$; see Table 8). This suggests that, when present, the front yard would be the most efficient signal for one to assess, as it is closely predictive of both general disorder and social disorganization.

We used the PCA results to create two new address descriptors by summing those variables that loaded on each component. Two addresses did not contain driveways in any of the four images but did contain streets and sidewalks. To create scores for pavement quality, we ran a regression for the other 18 addresses predicting the relationship between these scores with and without the rating of the driveway included. This linear equation produced the estimated pavement scores for the other two.

Of the three variables derived from images of a neighborhood, house care correlated positively ($r = .59$, $p < .01$) and the presence of a business negatively ($r = -.56$, $p < .05$) with a neighborhood's social quality. There was no significant relationship between the social quality and the pavement ($r = .28$, *ns*). When all three variables derived from images were entered into a regression predicting the social quality, house care was the primary significant predictor ($B = 0.44$, $p < .05$), and the presence of a business only approached significance ($B = -0.39$, $p < .10$).

Table 8
Correlations Between Physical Features of Images and Neighborhood Descriptors at Image and Neighborhood Levels

Variable	1	2	3	4	5	6	7	8	9	10	11
1. Paint	*(93/20)*	.32 (18)	.80*** (20)	.64** (20)	.13 (20)	.02 (20)	.02 (20)	.65** (20)	−.64** (20)	.50* (20)	.25 (20)
2. Driveway	.41** (59)	*(60/18)*	.58* (18)	.65** (18)	.24 (18)	.38 (18)	.43 (18)	.44† (18)	.02 (18)	.58* (18)	.41† (18)
3. Grass	.52*** (89)	.78*** (58)	*(95/20)*	.84*** (18)	.33 (20)	.23 (20)	.28 (20)	.66** (20)	−.54* (20)	.73*** (20)	.45* (20)
4. Vegetation	.56*** (92)	.69*** (60)	.76*** (95)	*(99/20)*	.39† (20)	.40† (20)	.31 (20)	.73*** (20)	−.16 (20)	.43† (20)	.36 (20)
5. Garbage	.38*** (93)	.44** (60)	.45*** (95)	.43*** (99)	*(100/20)*	.08 (20)	.40 (20)	.42† (20)	.14 (20)	.21 (20)	.11 (20)
6. Sidewalk	.17 (88)	.44** (58)	.38*** (88)	.37*** (92)	.15 (93)	*(93/20)*	.38 (20)	.27 (20)	.30 (20)	.02 (20)	−.13 (20)
7. Streets	.13 (86)	.33* (57)	.26* (88)	.24* (92)	.25* (93)	.37*** (90)	*(93/20)*	.05 (20)	.22 (20)	.27 (20)	−.15 (20)
8. Decorations	.17 (63)	.41** (39)	.31* (63)	.49*** (64)	.33** (65)	.21 (59)	.11 (58)	*(65/20)*	−.27 (20)	.48* (20)	.41† (20)
9. Business[a]	−.32** (93)	.06 (60)	.02 (95)	−.01 (99)	.02 (100)	.21* (93)	.15 (93)	−.09 (65)	*(100/20)*	−.56* (20)	−.34 (20)
10. Social quality										—	.55* (20)
11. Median income											—

Note. Neighborhood-level variables appear above the diagonal (shown in italics), and image-level variables appear below. Correlations for neighborhood descriptors are reported above the diagonal; correlations below the diagonal were used in the principal component analysis. The diagonal contains number of images or neighborhoods that include a score for each item. Correlations note number of cases with scores for both variables in parentheses. Names of ratings of physical features are simplified. See Table 7 for more information.

† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 9
Results of Principal Component Analysis, Including Variable Loadings and Eigenvalues for All Components

Variable	I	II	Communality
Paint	.72		.54
Driveway		.60	.59
Grass	.71		.71
Vegetation	.77		.74
Garbage	.68		.46
Sidewalk		.80	.64
Streets		.76	.61
Decorations	.71		.52
Eigenvalue	3.68	1.11	

Note. Only loadings above .5 are reported ($N = 80$).

When a stepwise regression was run with all physical measurements listed in Table 8, the grass quality was the only significant predictor ($B = 0.92$, $p < .01$; cut-point, $\alpha = .05$).

Disorder, disorganization, and attitudes. In Table 10 we report two models that use features of the images, neighborhood social quality (DAP), and an individual's knowledge of Binghamton to predict the responses seen in Study 1. In order of effect size, Model 1 found the maintenance of private houses and lawns ($B = 4.80$, $d = 0.78$, $p < .001$), the absence of businesses ($B = -5.87$, $d = 0.47$, $p < .01$), and the maintenance of paved surfaces ($B = 2.28$, $d = 0.39$, $p < .05$) to give raters a positive impression of an unfamiliar neighborhood, accounting for nearly all (89%) of the variation between neighborhoods. These factors fully mediate the relationship between observer and resident ratings of a neighborhood, implying that these elements are what the participants used to inform their judgments. Although we mentioned above that a quick judgment of the front lawn might suffice for an accurate impression of the neighborhood, this variable was not significantly associated with ratings when introduced in Model 2.

Disorder, disorganization, and behavior. When attempting to determine the physical cues informing behavior in the SPD, the first model used above was recast as a logit model, as the outcome variables are dichotomous. Because of the extreme collinearity between a neighborhood's social quality (DAP) and house care in the nine neighborhoods used in Study 2 ($r = .92$, $p < .001$), the former was excluded from analyses, as we were primarily interested in those cues in the images that inform behavior. None of those variables—house care, the presence of a business, or pavement care—significantly predicted one's decisions as a first mover, and only the presence of a business was associated with less reciprocation ($B = -0.26$, $p < .05$, $OR = 0.77$; see Table 11). The relationship between one's knowledge of Binghamton and reciprocity was unchanged in the new analyses.

Discussion

The first two studies provided evidence for community perception as a cognitive mechanism that effectively judges the safety of neighborhoods and adjusts social behavior accordingly. Here we found some evidence that individuals base these responses on a thorough observation of the neighborhood's physical structures, including those whose upkeep did not actually correlate with social quality (e.g., the maintenance of paved surfaces). This was clear in the reanalysis of the data from Study 1, and the null results in the reanalysis of the behavior in the SPD may also support this interpretation, as no specific feature was able to significantly predict cooperative behavior. Although this is a limitation of the small number of neighborhoods being compared ($N = 9$), it also suggests that no specific feature was predominant in influencing behavior.

Despite being the most informative feature in reality, the quality of the front lawn was not primarily associated with responses. Although this seems nonadaptive, the same correlation might not be consistent across cities or cultures—particularly those in which lawns are uncommon—thus community perception would be unlikely to have evolved with such a narrow orientation. Instead, it appears that observers responded to an array of elements, each reflecting the level of effort invested in the neighborhood's upkeep. An alternative hypothesis to this tendency for generalized assessment, though not mutually exclusive, is that individuals differ in the cues they use to judge a neighborhood. Such differences may be localized in individuals or correlated with one's cultural history. Regarding the negative reaction to the presence of businesses, it may be that participants perceived them to be a detriment to the development of a healthy community or, more simply, had an aversion to living beside businesses themselves. Further research will be necessary to answer these sorts of questions.

General Discussion

A previous study in the field of urban planning demonstrated that residents of a city consistently favor those neighborhoods whose structures appear orderly and well maintained (Nasar, 1990). Moreover, criminologists working in a variety of cities have repeatedly shown that these features correlate strongly with a neighborhood's social governance and safety (Corman & Mocan,

Table 10
Parameter Estimates From Multilevel Models Using the Physical Features of Images to Predict Ratings of Neighborhood Social Quality

	Model 1		Model 2	
Variable	*B*	*SE*	*B*	*SE*
Individual-level predictor				
Knowledge of Binghamton	1.27***	0.33	1.26***	0.33
Neighborhood-level predictors				
Social quality[a]	0.10	0.23	0.05	0.25
House	4.80***	0.87	3.99*	1.75
Pavement	2.28*	0.97	2.09†	1.05
Business	−5.87**	1.69	−5.48*	1.93
Grass			4.28	8.13
Cross-level interactions				
House × Knowledge	−0.29	0.40	−0.27	0.46
Pavement × Knowledge	0.30	0.23	−0.49	0.27
Business × Knowledge	−0.31	0.24	0.13	0.50
Grass × Knowledge			2.90	2.04
Approximate first-level R^2	.01		.01	
Approximate second-level R^2	.89		.89	

Note. $N = 2{,}791$ nested in 20 neighborhoods.
[a] Accessed from the 2006 application of the Developmental Assets Profile.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 11
Parameter Estimates From Multilevel Logit Models Using Physical Features of Images to Predict the Likelihood of Cooperating When Playing the Sequential Prisoner's Dilemma With a Resident of a Pictured Neighborhood

	Offers of cooperation (trust)			Reciprocation		
Variable	*B*	*SE*	*OR*	*B*	*SE*	*OR*
Individual-level predictors						
Knowledge of Binghamton	−0.04	0.10	0.96	0.32**	0.10	1.38
Age	0.15	0.09	1.16	0.02	0.10	1.02
Sex[a]	0.03	0.20	1.03	0.09	0.24	1.09
Neighborhood-level predictors						
House	0.03	0.08	1.03	−0.02	0.04	0.98
Pavement	0.20†	0.09	1.22	0.10	0.08	1.11
Business	−0.12	0.13	0.89	−0.26*	0.07	0.77

Note. $N = 294$ nested in nine neighborhoods. *OR* = odds ratio.
[a] Female = 1.
† $p < .10$. * $p < .05$. ** $p < .01$.

2005; Harcourt & Ludwig, 2006; Kawachi et al., 1999; Kelling & Sousa, 2001; Markowitz et al., 2001; Ross & Jang, 2000; Ross et al., 2001; Sampson & Raudenbush, 1999). Here we replicated both of these findings and, at their intersection, provide evidence that our preferences for well-maintained neighborhoods are not merely aesthetic but also serve a social function. When observing an urban landscape, participants formulated judgments about the quality of the local social environment that were highly accurate. These judgments also produced behaviors that were appropriate for these social expectations. The information encoded in disorder, however, was primarily useful in creating impressions of the local community and not necessarily of individual residents. Judgments of well-being, healthy habits, and prosociality were all inaccurate, suggesting that the attention to disorder serves the explicit purpose of assessing a community's social environment.

Taken together, these results provide evidence for an adaptation specific to group living, one that would have had to evolve in the short time that humans have lived in large settlements. We suggest that the most parsimonious way for this to have occurred is by merging preexisting mechanisms for assessing environments and social partners. Evidence for or against this hypothesis would require further research; studies of the more basic biological mechanisms underlying community perception would be particularly valuable. In addition, new studies should go beyond the species-specific approach taken here by exploring individual and cross-cultural differences in assessing neighborhoods.

The results also have strong applications for the field of sociology, where we feel it will be useful in defining the interplay between two major theories regarding the causes of crime. One proposes that it is the disorder of a neighborhood that invites crime (disorder theory; J. Q. Wilson & Kelling, 1982), the other that a lack of social governance permits crime to occur (social disorganization theory; Shaw & McKay, 1969). Although there is greater empirical support for the latter, many survey studies have found that disorder still weighs heavy in the minds of residents, influencing their perception of crime and the local community (Markowitz et al., 2001; Ross & Jang, 2000; Ross et al., 2001; Sampson & Raudenbush, 1999). Here we find that naive observers also respond to the apparent maintenance of an urban landscape, leading to attitudes and behaviors that are appropriate to the local social environment. Under the assumption that this same process is ongoing in residents of a deteriorating neighborhood, it comes as no surprise that researchers have found that individuals who live in such an environment tend to have an exaggerated sense of mistrust and helplessness (Ross & Jang, 2000; Ross et al., 2001). Thus, although the relationships between neighbors are most responsible for the actual crime rate, our attunement to the level of disorder may be influential in the development of these relationships. In turn, disorder itself becomes a highly relevant factor when considering how our built environments influence the attitudes and behaviors of residents and passersby alike.

References

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgment accuracy from thin slices of the behavioral stream. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 32, pp. 201–271). San Diego, CA: Academic Press. doi:10.1016/S0065-2601(00)80006-4

Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology, 86,* 599–614. doi:10.1037/0022-3514.86.4.599

Bowles, S., & Gintis, H. (2004). The evolution of strong reciprocity: Cooperation in heterogeneous populations. *Theoretical Population Biology, 65,* 17–28. doi:10.1016/j.tpb.2003.07.001

Brosig, J. (2002). Identifying cooperative behavior: Some experimental results in a prisoner's dilemma game. *Journal of Economic Behavior and Organization, 47,* 275–290. doi:10.1016/S0167-2681(01)00211-6

Corman, H., & Mocan, N. (2005). Carrots, sticks, and broken windows. *Journal of Law and Economics, 48,* 235–266. doi:10.1086/425594

Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience, 19,* 1508–1519. doi:10.1162/jocn.2007.19.9.1508

Fetchenhauer, D., Groothuis, T., & Pradel, J. (2010). Not only states but traits—Humans can identify permanent altruistic dispositions in 20 s. *Evolution and Human Behavior, 31,* 80–86. doi:10.1016/j.evolhumbehav.2009.06.009

Frijda, N. H. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology, 57,* 212–228. doi:10.1037/0022-3514.57.2.212

Gallup, A. C., & Wilson, D. S. (2009). Body mass index (BMI) and peer aggression in adolescent females: An evolutionary perspective. *Journal of Social, Evolutionary, and Cultural Psychology, 3,* 356–371.

Gangestad, S. W., Simpson, J. A., DiGeronimo, K., & Biek, M. (1992). Differential accuracy in person perception across traits: Examination of a functional hypothesis. *Journal of Personality and Social Psychology, 62,* 688–698.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology, 206,* 169–176.

Guadagnoli, E., & Velicer, W. F. (1988). Relation to sample size to the stability of component patterns. *Psychological Bulletin, 103,* 265–275. doi:10.1037/0033-2909.103.2.265

Harcourt, B. E., & Ludwig, J. (2006). Broken windows: New evidence from New York City and a five-city social experiment. *University of Chicago Law Review, 73,* 271–320.

Kaplan, S. (1992). Environmental preference in a knowledge-seeking, knowledge-using organism. In J. H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 581–598). New York, NY: Oxford University Press.

Kawachi, I., Kennedy, B. P., & Wilkinson, R. G. (1999). Crime: Social disorganization and relative deprivation. *Social Science and Medicine, 48,* 719–731. doi:10.1016/S0277-9536(98)00400-6

Kelling, G. L., & Sousa, W. H., Jr. (2001). *Do police matter? An analysis of the impact of New York City's police reforms* (Civic Report No. 22). New York, NY: Center for Civic Innovation at the Manhattan Institute.

Lazarus, R. S. (1991). *Emotion and adaptation.* New York, NY: Oxford University Press.

Levenson, R. W. (1994). Human emotions: A functional view. In P. Ekman & R. J. Davidson (Eds.), *The nature of emotion: Fundamental questions* (pp. 123–126). New York, NY: Oxford University Press.

Markowitz, F. E., Bellair, P. E., Liska, A. E., & Liu, J. (2001). Extending social disorganization theory: Modeling the relationships between cohesion, disorder, and fear. *Criminology, 39,* 293–319. doi:10.1111/j.1745-9125.2001.tb00924.x

Milgram, S., Mann, L., & Harter, S. (1965). The lost-letter technique: A tool of social research. *Public Opinion Quarterly, 29,* 437–438. doi:10.1086/267344

Nasar, J. L. (1990). The evaluative image of the city. *Journal of the American Planning Association, 56,* 41–53. doi:10.1080/01944369008975742

O'Brien, D. T., Wilson, D. S., Eldakar, O. T., & Carpenter, J. P. (2007). [Experimental economics at Binghamton High School]. Unpublished raw data.

Oltmanns, T. F., Friedman, J. N. W., Fiedler, E. R., & Turkheimer, E. (2004). Perceptions of people with personality disorders based on thin slices of behavior. *Journal of Research in Personality, 38,* 216–229. doi:10.1016/S0092-6566(03)00066-7

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, USA, 105,* 11087–11092. doi:10.1073/pnas.0805664105

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. T., Jr. (2004). *HLM 6: Hierarchical linear and nonlinear modeling.* Lincolnwood, IL: Scientific Software International.

Ross, C. E., & Jang, S. J. (2000). Neighborhood disorder, fear, and mistrust: The buffering role of social ties with neighbors. *American Journal of Community Psychology, 28,* 401–420. doi:10.1023/A:1005137713332

Ross, C. E., Mirowsky, J., & Pribesh, S. (2001). Powerlessness and the amplification of threat: Neighborhood disadvantage, disorder, and mistrust. *American Sociological Review, 66,* 568–591. doi:10.2307/3088923

Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology, 105,* 603–651. doi:10.1086/210356

Sampson, R. J., Raudenbush, S. W., & Earls, F. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science, 277,* 918–924. doi:10.1126/science.277.5328.918

Sapnas, K. G., & Zeller, R. A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement, 10,* 135–154. doi:10.1891/jnum.10.2.135.52552

Shaw, C. R., & McKay, H. D. (1969). *Juvenile delinquency and urban areas: A study of rates of delinquency in relation to differential characteristics of local communities in American cities* (Rev. ed.). Chicago, IL: University of Chicago Press.

Steblay, N. M. (1987). Helping behavior in urban and rural environments: A meta-analysis. *Psychological Bulletin, 102,* 346–356. doi:10.1037/0033-2909.102.3.346

Tooby, J., & Cosmides, L. (1990). The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and Sociobiology, 11,* 375–424. doi:10.1016/0162-3095(90)90017-Z

Verplaetse, J., Vanneste, S., & Braeckman, J. (2007). You can judge a book by its cover: The sequel. A kernel of truth in predictive cheating detection. *Evolution and Human Behavior, 28,* 260–271. doi:10.1016/j.evolhumbehav.2007.04.006

Wilson, D. S., O'Brien, D. T., & Sesma, A. (2009). Human prosociality from an evolutionary perspective: Variation and correlations at a city-wide scale. *Evolution and Human Behavior, 30,* 190–200. doi:10.1016/j.evolhumbehav.2008.12.002

Wilson, D. S., & Wilson, E. O. (2007). Rethinking the theoretical foundation of sociobiology. *Quarterly Review of Biology, 82,* 327–348. doi:10.1086/522809

Wilson, J. Q., & Kelling, G. L. (1982). The police and neighborhood safety: Broken windows. *Atlantic Monthly, 127,* 29–38.

(*Appendix follows*)

Appendix

Items Composing Scales Rated by Observers

Social cohesion
People around here are willing to help their neighbors.
There are adults in this neighborhood that children can look up to.
Social control
This is a safe neighborhood.
If there were a fight in this neighborhood, neighbors would interfere.
If children in this neighborhood were skipping school and hanging out on a street corner, neighbors would take action.
Well-being[a]
I feel good about the future.
Healthy habits[a]
I avoid things that are dangerous or unhealthy.
Prosociality[a]
I am sensitive to the needs and feelings of others.

[a] Preceded by the phrase "How do you think an adolescent living in this neighborhood would answer the following questions?"

Received November 20, 2009
Revision received November 13, 2010
Accepted November 18, 2010 ■