

Disagreeableness as a Cause and Consequence of Ostracism

Andrew H. Hales¹, Matthew P. Kassner², Kipling D. Williams¹,
and William G. Graziano¹

Personality and Social
Psychology Bulletin
2016, Vol. 42(6) 782–797
© 2016 by the Society for Personality
and Social Psychology, Inc
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0146167216643933
pspb.sagepub.com



Abstract

Ostracism's negative consequences have been widely documented, but research has yet to explore the personality characteristics of its targets that precipitate ostracism. Based on theories of the functions of ostracism, we found that people are more willing to ostracize disagreeable targets than more agreeable targets (Studies 2 and 3). This outcome was mediated by participants' interpersonal trust toward the target, and was especially strong for people who highly endorse fairness as a foundation for morality (Study 4). Ironically, the experience of ostracism induced a state of disagreeableness: the very characteristic that elicits ostracism from others (Study 5). This relationship was mediated by feelings of anger (Study 6). Findings indicate disagreeableness is a particularly negative outcome of ostracism, because it leads to further ostracism.

Keywords

ostracism, agreeableness, personality, fairness moral foundation

Received December 13, 2014; revision accepted March 20, 2016

People differ widely in agreeableness. Some people tend to be kind, pleasant, and trusting; others are unkind, unpleasant, and untrusting. These differences are surprisingly consequential, with disagreeable people suffering negative life outcomes including reduced longevity, job attainment, and psychological health (Ozer & Benet-Martinez, 2006). In this article, we explore the possibility that people who show cues of disagreeableness, aside from experiencing these negative outcomes, also suffer from ostracism more frequently than those who do not show cues of disagreeableness. This is important, given that ostracism itself is associated with a wide range of negative consequences including reduced belonging, self-esteem, control, and meaningful existence (Williams, 2009). Ostracism also leads to antisocial behaviors, including aggression (Williams & Wesselmann, 2011). In this article, we also explore the complimentary possibility that ostracism leads not only to antisocial behaviors but also to a generally disagreeable state.

We propose that ostracism and agreeableness are negatively bidirectionally related; disagreeableness elicits ostracism from others, which, in turn, leads to low agreeableness. In this article, we consider these two distinct hypotheses. We will first address the rationale for disagreeable people being more likely to receive ostracism, and then turn our attention to why ostracism might induce a state of disagreeableness.

Disagreeableness as a Cause of Ostracism

Much is known about the consequences of ostracism, but its antecedents have received less attention. Early theorizing outlined various reasons why a source might ostracize a target (Williams, 1997). For example, *punitive ostracism* involves using ignoring and excluding as tools to achieve retribution for a perceived wrongdoing (e.g., the silent treatment). Another noted motive is managing behaviors that are inappropriate, deviant, or burdensome (Gruter & Masters, 1986; Wesselmann, Wirth, Pryor, Reeder, & Williams, 2013). This literature has emphasized behavioral antecedents of ostracism, leaving unexamined the possibility that there are certain personality characteristics that render some people especially likely targets.

Consistent with this possibility, theorists have proposed that humans have evolved a set of behavioral adaptations to socially exclude individuals who are poor social exchange

¹Purdue University, West Lafayette, IN, USA

²Centre College, Danville, KY, USA

Corresponding Author:

Andrew H. Hales, Purdue University, 703 Third Street, West Lafayette, IN 47907, USA.

Email: ahales@purdue.edu

partners (Gruter & Masters, 1986; Kurzban & Leary, 2001). Specifically, Kurzban and Leary (2001) argue that, in addition to minimizing pathogen threat, social exclusion functions to (a) help *individuals* avoid being taken advantage of in dyadic interactions, and (b) help *groups* enhance fitness relative to other groups. Both of these functions require that interaction partners or group members be at least somewhat cooperative. Accordingly, research shows that people are more cooperative in social dilemmas when threatened with the possibility of ostracism (Feinberg, Willer, & Schultz, 2014), and also that people can be more willing to fight and die for an important ingroup after being ostracized (Gómez, Morales, Hart, Vázquez, & Swann, 2011).

Given the relationship between agreeableness and cooperation, we propose that people showing cues of disagreeableness are especially likely to be ostracized. Agreeableness has been conceptualized as the motivation to maintain positive relationships (Graziano & Eisenberg, 1997). Research attests to the power of agreeableness in achieving this end (see Graziano & Tobin, 2009, for a review). Disagreeable people have been documented to be less empathic, and less helpful (e.g., Graziano, Habashi, Sheese, & Tobin, 2007). They also defect more in prisoner's dilemma scenarios (Kagel & McGee, 2014). These findings suggest that disagreeable people are less trustworthy interaction partners. Given the evolutionary functions of social exclusion, we hypothesized that disagreeable individuals will be likely targets of ostracism.

One related study explored the relationship between agreeableness and social interactions in middle schoolchildren. Disagreeable children have fewer friends overall (Study 1). Moreover, in a longitudinal study, disagreeable children were bullied more if they displayed certain behavioral predispositions (e.g., poor social skills). Conversely, agreeable children could display the same behavioral predispositions without increased bullying (Study 2; Jensen-Campbell, Adams, et al., 2002). This research underscores the importance of agreeableness in securing social inclusion.

We hypothesized that people are most likely to ostracize targets who are disagreeable. To test this hypothesis, we examined ostracism from the sources' perspective, looking at their responses to targets varying in agreeableness.

Disagreeableness as a Consequence of Ostracism

We turn now to the distinct, but related question of whether being ostracized induces a state of disagreeableness. Support for this hypothesis comes from research documenting the antisocial behavioral consequences of ostracism. For example, unless participants were given an opportunity to restore control, following a brief ostracism experience, they allocated more hot sauce to a confederate believing they would have to eat it all (Warburton, Williams, & Cairns, 2006). Likewise, participants who were rejected by a group subsequently assigned more hot sauce to an innocent interaction partner,

especially when the rejection was unexpected (Wesselmann et al., 2010). Similar findings have been observed using a variety of rejection manipulations and aggression measures (see Williams & Wesselmann, 2011, for a review).

In addition to instigating antisocial behavior, social rejection reduces empathy, a major component of agreeableness (Graziano et al., 2007). Reduced empathy leads to decreases in a wide range of prosocial behaviors including cooperation in a prisoner's dilemma (Twenge, Baumeister, DeWall, Ciarocco, & Bartels, 2007). Social rejection has dual effects of increasing antisocial behavior and decreasing prosocial behavior, behavior patterns that are characteristic of disagreeable people. Based on this observation, we hypothesize that ostracism would induce a state of disagreeableness.

We acknowledge that people are not uniformly antisocial after being ostracized. A similarly large literature finds that ostracism can motivate behaviors that are deferential in nature and help individuals achieve reinclusion in groups and reformation of threatened needs (Smart Richman & Leary, 2009; Williams, 2009). For example, Riva, Williams, Torstrick, and Montali (2014) note that ostracism has been documented to increase vulnerability to the three major forms of social influence: conformity, compliance, and obedience. The current research offers insight into when people respond to ostracism prosocially versus antisocially. If ostracism induces a disagreeable disposition, it would raise the possibility that prosocial responses following ostracism represent *outward displays* intended to fortify threatened needs (Williams, 2009). These outward displays would occur despite an inwardly experienced state of disagreeableness.

A secondary purpose of this article is to test potential mediators of the effect of ostracism on agreeableness. We tested several known outcomes of ostracism as potential mediators of the hypothesized effect of ostracism on disagreeableness. Ostracism threatens four basic needs: belonging, self-esteem, control, and meaningful existence. It also has a strong effect on mood, decreasing positive affect and increasing negative affect. Given that agreeableness is reliably associated with positive affect and emotional stability (John & Srivastava, 1999), it is plausible that negative affect produced by ostracism leads to a state of low agreeableness.

Rather than only examining the valence of mood, we considered which specific emotions might account for changes in agreeableness (Lerner & Keltner, 2000). We were particularly interested in the differential roles of sadness and anger in producing disagreeableness. Anger is characterized as an activating emotion highly tied to aggression (Berkowitz & Harmon-Jones, 2004). Although sadness is also subjectively unpleasant to experience, it does not have the same implications for antisocial tendencies. In line with this reasoning, Chow, Tiedens, and Govan (2008) found that ostracism induced both sadness and anger, but only anger predicted subsequent antisocial behavior. Accordingly, we tested anger and sadness independently as potential mediators of the effect of ostracism on state disagreeableness.

Research Overview

After verifying that agreeableness is associated with lower levels of chronic ostracism (Study 1), we experimentally tested whether disagreeableness elicits ostracism in two experiments in which participants reported their willingness to ostracize targets varying in agreeableness. Study 2 tested whether a social violation is necessary for disagreeableness to elicit ostracism. Studies 3 and 4 tested the possibility that people *include* agreeable targets rather than *ostracize* disagreeable targets. In addition, Study 4 tested interpersonal trust as a mediator of the effect of agreeableness on intentions to ostracize, and moral concerns with fairness as a moderator of that mediation.

To test the reverse causal direction, we conducted two experiments in which the participants themselves were targets of ostracism or inclusion. They then reported their state level of agreeableness. In the final experiment, we also measured need satisfaction, anger, and sadness as potential mediators of the effect of ostracism on agreeableness.

Study 1

The goal of Study 1 was to assess whether disagreeableness is correlated with chronic ostracism experiences.

Method

Students ($N = 792$; 394 males) took an online questionnaire for partial credit in an introductory psychology course. Participants' average age was 19.60 ($SD = 1.75$) years. Measures of agreeableness and ostracism experiences were embedded in a series of unrelated questionnaires that were part of an initial prescreening survey administered at the beginning of the semester. The measures were presented in random order.

Participants reported their agreeableness by responding to nine items from the Big Five Inventory (e.g., "I see myself as someone who is helpful and unselfish with others"; John, Donahue, & Kentle, 1991; $\alpha = .76$). Ratings were made on a scale from 1 (*strongly disagree*) to 5 (*strongly agree*). Participants also reported their experiences with ostracism by completing the eight-item Ostracism Experiences Scale (e.g., "In general, others treat me as if I am invisible"; Carter-Sowell, 2010; $\alpha = .94$). Ratings were made on a scale from 1 (*hardly ever*) to 7 (*almost always*).

Results and Discussion

Agreeableness was negatively correlated with ostracism experiences, $r(790) = -.34$, $p < .001$, 95% confidence interval (CI) = $[-0.40, -0.28]$. This relationship was not reduced after controlling for age and gender, partial $r(788) = -.34$.

The finding that agreeableness is related to ostracism experiences is consistent with our prediction that agreeableness and

ostracism are bidirectionally causally related, but is limited by its correlational design. In Study 2, we experimentally tested whether people would be more likely to ostracize someone described as low in agreeableness than someone described as high in agreeableness. We also tested whether a disagreeable person must commit a social infraction to provoke ostracism.

Study 2

Method

Participants and design. Introductory psychology students ($N = 113$) completed an online experiment of impressions of others for partial course credit (76 males, $M_{Age} = 20.55$, $SD_{Age} = 2.04$). Participants were randomly assigned to read and respond to one of four vignettes in a 2 (target agreeableness: agreeable, disagreeable) \times 2 (target behavior: prosocial, antisocial) between-subjects design.

Procedure. Participants were presented with a short vignette describing a target individual, *Mason*. They were instructed to imagine what type of person Mason is, and what it would be like to spend time around him. The vignette (210 words) described a 19-year-old who majors in engineering and enjoys cooking. Those in the high agreeable conditions read within the vignette that, "Mason tends to be a warm, trusting, and caring person." Those in the low agreeable conditions instead read that, "Mason tends to be a cold, untrusting, and uncaring person." These descriptions were based on items from the Big Five Inventory (John et al., 1991).

After reading a description of the target's interests and personality, participants were asked to rate their agreement with three statements, ostensibly to verify passage comprehension. Two of the statements were irrelevant filler items. The third statement, "Mason is agreeable," served as the target personality manipulation check.

Following this manipulation check, participants read the remaining portion of the vignette (presented on the next page of the online survey), which described Mason's job and also how Mason responded when his friend, Laura, asked him to help her move furniture on a Saturday. Those in the prosocial conditions read that, "Mason had planned to spend the day hiking, but he still agreed to help Laura move. Laura and Mason spent the entire day moving furniture." Those in the antisocial conditions read that, "Mason had planned to spend the day hiking, so he refused to help Laura move. Laura spent the day moving furniture without Mason."

Following the vignette, participants were presented with another two statements to verify passage comprehension. One statement was an irrelevant filler item. The other statement, "Mason agreed to help Laura move" was a check of the target behavior manipulation. Following the manipulation check, participants completed questionnaires assessing intentions to confront and ostracize Mason, how much they

Table 1. Main Effects of Target Personality and Target Behavior on Outcomes in Study 2 ($n = 113$).

	Agreeable vs. disagreeable target				Prosocial vs. antisocial target			
	Agreeable <i>M</i> (<i>SD</i>)	Disagreeable <i>M</i> (<i>SD</i>)	<i>d</i>	95% CI	Prosocial <i>M</i> (<i>SD</i>)	Antisocial <i>M</i> (<i>SD</i>)	<i>d</i>	95% CI
Ostracism intentions	1.86 (0.78)	2.57 (0.75)	0.93	[0.56, 1.32]	2.05 (0.82)	2.36 (0.84)	0.36	[-0.01, 0.73]
Confrontation intentions	2.16 (0.85)	2.56 (1.11)	0.40	[0.03, 0.77]	1.95 (0.92)	2.79 (0.91)	0.92	[0.53, 1.31]
Liking	3.69 (0.77)	2.53 (0.76)	1.51	[1.09, 1.93]	3.48 (0.84)	2.76 (0.94)	0.81	[0.42, 1.19]
Belief target's choice was wrong	2.18 (1.34)	2.67 (1.52)	0.34	[-0.04, 0.71]	1.60 (0.99)	3.27 (1.34)	1.39	[0.99, 1.81]
Surprised by target's choice	2.63 (1.26)	3.24 (1.26)	0.48	[0.11, 0.86]	2.70 (1.35)	3.16 (1.20)	0.36	[-0.01, 0.73]
Impressed by target's choice	2.89 (1.32)	2.87 (1.45)	-0.01	[-0.39, 0.36]	3.75 (1.12)	2.00 (1.02)	-1.63	[-2.02, -1.21]
Pleased by target's choice	3.23(1.17)	2.89 (1.51)	-0.25	[-0.63, 0.12]	4.00 (0.96)	2.09 (0.95)	-1.97	[-2.44, -1.53]

Note. Responses for all items are on a 5-point scale. CI = confidence interval.

like Mason, and finally, individual differences including ostracism experiences and all Big Five personality traits.

Measures. Intentions to ostracize the target were measured by asking participants to imagine that Mason just joined a campus club that they are in and to rate their agreement with seven items assessing the likelihood of engaging in behaviors such as excluding, ignoring, and giving the silent treatment (e.g., “I might find myself ignoring Mason”; $\alpha = .89$; see online reporting for full scale). Next, participants reported intentions to confront Mason so we could verify that any effect of target personality on ostracism intentions reflected a desire to avoid the target above and beyond a desire to express disapproval or disliking. Intentions to confront the target were measured by asking participants to rate their agreement with three statements in reference to Mason's decision either to help, or not help Laura move (e.g., “I would confront Mason directly about what he did”; $\alpha = .79$). Next, they rated the extent to which they believed Mason's choice was wrong, and how surprised, impressed, and pleased they were with the decision (all separate items). Ratings were made on a scale from 1 (*completely disagree*) to 5 (*completely agree*). Next participants responded to the question, “If you had to choose only one way to respond to Mason's decision, which would you pick?” followed by the choices, “confront him,” “ostracize him,” and “neither.” Participants reported their liking toward Mason with five items (e.g., “I feel like Mason and I would get along”; $\alpha = .92$).

Results

Manipulation checks. A 2 (target personality) \times 2 (target behavior) ANOVA on the agreeableness manipulation check revealed only a significant main effect for personality condition $F(1, 109) = 120.14, p < .001, d = 2.08, 95\% \text{ CI} = [1.62, 2.54]$. Participants in the agreeable target condition rated Mason as more agreeable ($M = 4.38, SD = 0.88$) than those in

the disagreeable target condition ($M = 2.35, SD = 1.08$). Likewise, the same analysis conducted on the behavior manipulation check revealed only a significant main effect for behavior condition, $F(1, 109) = 489.85, p < .001, d = 4.15, 95\% \text{ CI} = [3.49, 4.81]$, with those in the prosocial behavior condition ($M = 4.65, SD = 0.88$) correctly reporting that he helped his friend more so than those in the antisocial behavior condition ($M = 1.27, SD = 0.75$).

Reactions to the target. Means and CIs for the dependent variables are provided in Table 1. We ran the same target personality by target behavior ANOVA on the average of the seven-item composite measure of intentions to ostracize the target. Results revealed a significant main effect for target personality, $F(1, 109) = 25.01, p < .001, d = 0.93, 95\% \text{ CI} = [0.54, 1.32]$. People rated higher willingness to ostracize Mason if he was described as disagreeable ($M = 2.57, SD = 0.75$) than if he was described as agreeable ($M = 1.86, SD = 0.78$). There was also a significant main effect for behavior, $F(1, 109) = 4.66, p = .033, d = 0.36, 95\% \text{ CI} = [-0.003, 0.74]$. Participants were more likely to ostracize Mason if he had refused to help his friend move ($M = 2.36, SD = 0.84$) than if he had agreed to do so ($M = 2.05, SD = 0.83$). Target disagreeableness did not interact with target behavior, $F(1, 109) = 0.62, ns$. People were more likely to ostracize the disagreeable target when he refused to help, $t(54) = 2.83, p = .007, d = 0.76, 95\% \text{ CI} = [0.21, 1.30]$, and also when he agreed to help, $t(55) = 4.33, p < .001, d = 1.15, 95\% \text{ CI} = [0.58, 1.71]$. There were no significant main effects of gender, and only a marginal interaction between gender and behavior, $F(1, 105) = 3.16, p = .078, \eta_p^2 = .03$.

We also conducted the same analysis with liking the target as a covariate. Results showed that the extent to which participants liked the target was a very strong predictor of ostracism intentions, $F(1, 108) = 88.68, p < .001, \eta_p^2 = .45$. Controlling for liking, the effect of target personality became non-significant, $F(1, 108) = 1.31, p = .26, \eta_p^2 = .01$.

Table 2. The Number of People in Each Condition in Study 2 Choosing to Ostracize the Target, Confront the Target, or Take No Action.

Condition	<i>n</i>	Ostracize	Confront	Neither	χ^2	<i>p</i>
Prosocial behavior						
Agreeable target	28	1	6	21	23.21	<.001
Disagreeable target	28	0	5	23	11.57	.001
Antisocial behavior						
Agreeable target	28	1	11	16	12.50	.002
Disagreeable target	26	3	12	11	5.62	.06
Total	110	5	34	71	59.69	<.001

Note. Three participants did not provide a response to this question.

Finally, when forced to choose between confrontation, ostracism, or neither, participants generally avoided ostracism (Table 2). Chi-square independence tests showed that the target behavior manipulation significantly increased the proportion of participants choosing to confront the target, $\chi^2(2) = 10.07, p = .006$. The target personality manipulation did not affect this decision, $\chi^2(2) = 0.20, p = .865$. To test the interaction, we collapsed the confrontation and ostracism response options into one category and performed a logistic regression testing whether the manipulations interacted to predict the likelihood of selecting some action versus selecting neither. The interaction was not significant, $b = 1.03, z = 1.19, 95\% \text{ CI} = [-0.66, 2.71]$.

Discussion

Study 2 demonstrates that people are more willing to ostracize a target described as low in agreeableness relative to a target described as high in agreeableness. Overall, ostracism did not seem to be a popular response to the scenario; few people selected it in the forced choice question, and even in the disagreeable and antisocial condition, the mean score was barely above the scale midpoint. This is not surprising, considering that perpetrating ostracism is itself a psychologically uncomfortable experience (Legate, DeHaan, Weinstein, & Ryan, 2013). However, looking at the *relative* willingness to ostracize, we see a robust increase when the target is disagreeable.

Some ambiguity surrounds these results. Without a control condition, it is unclear whether people ostracize disagreeable targets, or whether they include agreeable targets. Study 3 adds clarity by including a control condition in which the target's personality is not described. Because agreeableness affected ostracism independently of prosocial versus antisocial behavior, we dropped this independent variable in Study 3.

The finding that the effect of target agreeableness became non-significant when controlling for liking suggests that disliking is a highly plausible candidate explanation for this effect. We continued to measure and account for liking in Study 3.

Study 3 also included an exploratory measure of the extent to which people are concerned about fairness when making moral decisions (Graham, Haidt, & Nosek, 2009). This allowed us to examine whether the effect of target personality on ostracism intentions is stronger for people who are highly concerned about fairness. Given that ostracism can serve the function of protecting groups from poor exchange partners (Kurzban & Leary, 2001), it follows that people who are especially concerned about the principle of justice will show the strongest tendency to ostracize disagreeable targets.

Study 3

Method

Participants and design. Introductory psychology students ($N = 122$) completed an online experiment of impressions of others for partial course credit (53 males, $M_{\text{Age}} = 19.12, SD = 1.18$). Participants were randomly assigned to one of three conditions: agreeable target, disagreeable target, or control (no personality information about target).

Procedure. Participants were asked to read and respond to a vignette that was identical to Study 2, except it did not include the description of Mason's friend asking him to move or his response.

The descriptions of Mason's personality in the high and low agreeableness conditions were the same as Study 2. In the control condition, the sentence describing Mason's personality was omitted. Ostracism intentions ($\alpha = .90$) and liking ($\alpha = .89$) were measured with the same items used in Study 2.

Participants also completed the 15-item moral foundations questionnaire used by Graham and colleagues (2009). The instructions read, "When you decide whether something is right or wrong, to what extent are the following considerations relevant to your thinking?" Fairness ($\alpha = .79$), harm ($\alpha = .78$), ingroup loyalty ($\alpha = .56$), authority ($\alpha = .35$), and sanctity ($\alpha = .68$) were each assessed with three items (e.g., fairness item "whether or not some people were treated

Table 3. Hierarchical Multiple Regression Analysis Predicting Ostracism Intentions From the Fairness Moral Foundation and Its Interaction With Target Agreeableness in Study 3.

Predictor	Model 1			Model 2		
	<i>b</i>	<i>t</i> -value	95% CI	<i>b</i>	<i>t</i> -value	95% CI
Step 1						
Fairness	-0.25**	-3.14	[-0.40, -0.09]	0.008	0.06	[-0.26, 0.27]
Dummy Code 1 (disagreeable vs. agreeable)	-0.98***	-6.82	[-1.26, -0.69]	-0.97***	-6.95	[-1.25, -0.70]
Dummy Code 2 (disagreeable vs. control)	-0.81***	-5.73	[-1.09, -0.53]	-0.80***	-5.73	[-1.07, -0.52]
Step 2						
Dummy Code 1 × Fairness				-0.50**	-2.71	[-0.86, -0.13]
Dummy Code 2 × Fairness				-0.24	-1.23	[-0.63, 0.15]
<i>R</i> ²	.34***	<i>F</i> = 20.45		.38***	<i>F</i> = 24.14	
<i>R</i> ² change				.04*	<i>F</i> = 3.69	

Note. Dummy Code 1 contrasts the agreeable and disagreeable conditions. Dummy Code 2 contrasts the control and disagreeable conditions. Betas are unstandardized, and fairness is mean centered. CI = confidence interval.

p* < .05. *p* < .01. ****p* < .001.

differently than others”). Responses were made on a scale from 1 (*never relevant*) to 5 (*always relevant*). This measure was included and analyzed on an exploratory basis.

Results

Manipulation check. A one-way ANOVA on ratings of agreeableness revealed significant differences between conditions, $F(2, 119) = 41.80, p < .001$. Tukey’s HSD (honest significant difference) test revealed that participants in the disagreeable target condition rated the target as less agreeable ($M = 2.44, SD = 0.94$) than those in the agreeable condition ($M = 4.33, SD = 1.01$), $t(119) = 8.26, p < .001, d = 1.94$, 95% CI = [1.41, 2.46], and the control condition ($M = 4.13, SD = 0.94$), $t(119) = 7.40, p < .001, d = 1.79$, 95% CI = [1.23, 2.30]. The agreeable condition and control condition did not significantly differ from each other, $t(119) = 0.89, p = .65, d = 0.20$, 95% CI = [-0.24, 0.64].

Ostracism intentions. Descriptions of the target’s personality affected intentions to ostracism him, $F(2, 119) = 23.95, p < .001, \eta_p^2 = .29$. Participants reported greater intentions to ostracize the disagreeable target ($M = 2.54, SD = 0.75$) than the agreeable target ($M = 1.60, SD = 0.64$), $t(119) = 6.39, p < .001, d = 1.36$, 95% CI = [0.86, 1.84], as well as the control target ($M = 1.75, SD = 0.61$), $t(119) = 5.39, p < .001, d = 1.16$, 95% CI = [0.69, 1.62]. Intentions to ostracize the agreeable target did not differ from the control target, $t(119) = 1.02, p = .57, d = 0.26$, 95% CI = [-0.18, 0.70]. Descriptions of the target’s personality also affected liking in a similar way, $F(2, 119) = 28.72, p < .001$. The disagreeable target ($M = 2.88, SD = 0.74$) was liked less than the agreeable target ($M = 3.87, SD = 0.67$), $t(119) = 6.59, p < .001, d = 1.41$, 95% CI = [0.92, 1.89], and also the control target ($M = 3.85, SD = 0.63$), $t(119) = 6.45, p < .001, d = 1.40$, 95% CI = [0.92,

1.88]. The agreeable and control targets did not differ, $t(119) = 0.19, p = .98, d = -0.05$, 95% CI = [-0.49, 0.39]. The effect of condition did not significantly interact with gender, $F(1, 116) = 2.19, p = .117, \eta_p^2 = .12$.

As in Study 2, liking was a strong predictor of ostracism intentions, $F(2, 118) = 70.89, p < .001, \eta_p^2 = .38$. However, unlike Study 2, the effect of agreeableness condition on intentions to ostracize remained significant even after controlling for liking, $F(2, 118) = 3.22, p < .001, \eta_p^2 = .05$.

Fairness concerns. To test whether the effect of target agreeableness on ostracism intentions was moderated by endorsement of the fairness foundation, we constructed two dummy coded variables with the disagreeable condition as a reference group. We created interaction terms by multiplying the dummy coded variables by endorsement of the fairness foundation (mean centered). When added together as a second step in a hierarchical regression, the interaction terms explained significant variance beyond the dummy codes and endorsement of the fairness foundation, which were entered in the first step, $F(2, 116) = 3.69, p = .028, \Delta R^2 = .04$ (see Table 3). Simple slopes tests reveal that fairness predicts reduced intentions to ostracize an agreeable target, $b = -0.49, t(116) = -3.89, p < .001$, and marginally reduced intentions to ostracize a control target, $b = -0.23, t(116) = -1.62, p = .11$, but not a disagreeable target, $b = 0.008, t(116) < 1$ (see Figure 1).

We also tested the simple effects of condition at low and high levels of fairness concerns. To be exhaustive, we also report the comparison between the agreeable and control target conditions. Participants low in fairness concerns (at 1 standard deviation below the mean) reported greater intentions to ostracize the disagreeable target than the control target, $b = -0.60, t(116) = -3.06, p = .003$, or agreeable target, $b = -0.62, t(116) = -3.02, p = .003$, which did not differ from

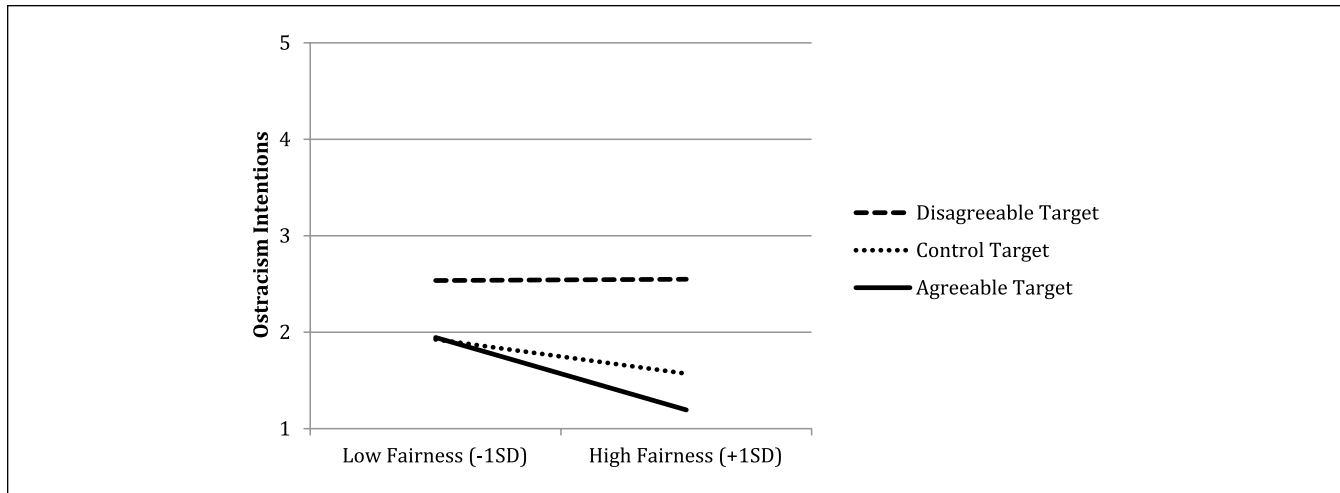


Figure 1. The relationship between endorsement of the fairness foundation and intentions to ostracize an agreeable, disagreeable, or control target in Study 3 ($n = 122$).

each other, $b = -0.02$, $t(116) = -0.08$, $p = .938$. Participants high in fairness concerns (at 1 standard deviation above the mean) also reported greater intentions to ostracize the disagreeable target than the control target, $b = -0.97$, $t(116) = -4.92$, $p < .001$, or agreeable target, $b = -0.60$, $t(116) = -1.34$, $p < .001$. However, participants high in fairness concerns reported marginally less intention to ostracize the agreeable target relative to the control target, $b = -0.37$, $t(116) = -1.80$, $p = .074$.

The interaction was not significant after controlling for liking, $F(2, 115) = 0.74$, $p = .48$, $\Delta R^2 = .005$.

Discussion

Study 3 helps clarify the causal relationship between agreeableness and ostracism. People were more likely to ostracize a disagreeable target relative not only to an agreeable target but also to a target for whom no personality information was provided. It would seem that disagreeableness elicits ostracism, rather than agreeableness eliciting inclusion. However, responses to the manipulation check introduce ambiguity into this conclusion; people regarded the control target as equally agreeable as the target who was explicitly described as agreeable. This suggests people expect others to be agreeable unless there is evidence to the contrary. On this interpretation, it remains unclear whether people ostracize those who are perceived as disagreeable or include those who are perceived as agreeable. Study 4 addresses this ambiguity by including a moderate agreeableness condition.

Unlike Study 2, the effect of agreeableness on intentions to ostracize remained significant even after controlling for liking toward the target. In the absence of any obvious methodological differences between studies that could explain this discrepancy, we continued to measure and account for liking. We discuss this variable further below.

Although it was not predicted a priori, we consider the observation that the moral foundation of fairness moderates the effect of agreeableness on ostracism intentions to be highly informative. The interaction shows that participants higher in fairness concerns are more discriminating between different levels of agreeableness, primarily because of reduced intentions to ostracize an agreeable target. The fairness moral foundation is thought to reflect humans' evolved careful attention to fairness, reciprocity, and protection against exploitation (Graham et al., 2009). Thus, this extra sensitivity to agreeableness in high-fairness participants suggests that the effect of agreeableness on ostracism intentions is partially motivated by a reduction in defensive concerns about being exploited when a target is agreeable, in addition to liking the target more. If so, then distrust toward the target should mediate the effect of agreeableness, especially for people who are high in fairness concerns. Study 4 tests this directly by measuring interpersonal trust. We predicted that disagreeable targets would be viewed as less trustworthy, which would increase ostracism intentions. Moreover, this mediational role of trustworthiness should be especially strong for participants who are high in fairness concerns.

We did not have a strong prediction for whether fairness would moderate the link between agreeableness and trust (which would suggest differential *detection* of trustworthiness based on one's fairness concerns) or, alternatively, the link between trust and ostracism intentions (which would suggest differential *weighting* of trustworthiness based on one's fairness concerns).

Study 4

Method

Participants and design. Introductory psychology students ($N = 273$) participated in an online study for partial course

Table 4. Means and Standard Deviations for Manipulation Check and Dependent Variables by Condition in Study 4 ($n = 273$).

	Condition		
	Agreeable target	Moderate target	Disagreeable target
Manipulation check	4.25 [4.04, 4.46] ^a	3.53 [3.36, 3.70] ^b	2.72 [2.48, 2.96] ^c
Interpersonal trust	3.80 [3.68, 3.93] ^a	3.57 [3.48, 3.67] ^b	2.98 [2.87, 3.10] ^c
Ostracism intentions	1.82 [1.69, 1.97] ^a	1.97 [1.83, 2.11] ^a	2.38 [2.22, 2.54] ^b
Liking	3.80 [3.67, 3.94] ^a	3.42 [3.29, 3.56] ^b	2.82 [2.68, 2.95] ^c

Note. Responses for all items are on a 5-point scale. 95% confidence intervals appear in brackets, means not sharing a subscript are significantly different, Tukey's HSD (honest significant difference) test.

credit (146 males, 126 females, one did not report gender, $M_{Age} = 19.72$, $SD = 1.39$). Participants were randomly assigned to read about a disagreeable target, a moderately agreeable target, or a fully agreeable target.

Procedure. Prior to reading the vignettes, participants completed the moral foundations questionnaire used in Study 3 (fairness $\alpha = .68$).

Participants read vignettes nearly identical to Study 3. In the disagreeable condition, participants read that the target "tends to be cold, uncaring, and untrusting." In the moderate condition, participants read that the target "tends to be neither warm nor cold, neither trusting nor untrusting, and neither caring nor uncaring." In the fully agreeable condition, participants read that the target "tends to be warm, caring, and trusting." After reading the vignette, participants responded to the same manipulation check used in Study 2.

Next, participants reported interpersonal trust toward the target by rating their agreement with 14 items from the Overall Trust and Reliability subscales of the Specific Interpersonal Trust Scale (Johnson-George & Swap, 1982; $\alpha = .89$; for example, "I would expect Mason to play fair"). Then, participants completed the same measures used in Study 3: ostracism intentions ($\alpha = .87$), liking ($\alpha = .87$), ostracism experiences ($\alpha = .94$), and Big Five traits (smallest $\alpha = .79$).

Results

Manipulation check. Descriptions of the target's personality affected ratings of agreeableness, $F(2, 270) = 53.81$, $p < .001$. As intended, there was a stepwise increase in ratings of agreeableness between disagreeable and moderate conditions, $t(270) = 5.56$, $p < .001$, $d = 0.81$, 95% CI = [0.51, 1.11], and then between moderate and agreeable conditions, $t(270) = 4.45$, $p < .001$, $d = 0.80$, 95% CI = [0.50, 1.10]. Importantly, the effect size of the differences is nearly equivalent between disagreeable and moderate conditions, and moderate and agreeable conditions (see Table 4). In the

moderate condition, participants viewed the target as truly moderately agreeable, rather than assuming he was highly agreeable as in Study 3.

Ostracism intentions. The agreeableness manipulation affected willingness to ostracize the target, $F(2, 266) = 19.48$, $p < .001$, $\eta_p^2 = .13$. However, a significant gender interaction revealed that the nature of this effect was different for males and females, $F(2, 266) = 5.53$, $p = .004$, $\eta_p^2 = .04$.¹

For females, ostracism intention increased linearly with target disagreeableness. They rated more willingness to ostracize the moderate target ($M = 1.87$, $SD = 0.66$) than the agreeable target ($M = 1.51$, $SD = 0.54$), $t(266) = 2.44$, $p = .016$, $d = 0.60$, 95% CI = [0.16, 1.04], and also more willingness to ostracize the disagreeable target ($M = 2.47$, $SD = 0.83$) than the moderate target, $t(266) = 4.17$, $p < .001$, $d = 0.80$, 95% CI = [0.36, 1.23].

In contrast, males showed a pattern similar to Study 3, with roughly equal intentions to ostracize the agreeable target ($M = 2.03$, $SD = 0.62$) and moderate target ($M = 2.08$, $SD = 0.65$), $t(266) = 0.35$, $p = .730$, $d = 0.08$, 95% CI = [0.33, 0.49]. However, they reported significantly greater ostracism intention toward the disagreeable target ($M = 2.31$, $SD = 0.73$) relative to the agreeable target, $t(266) = 2.12$, $p = .035$, $d = 0.42$, 95% CI = [0.03, 0.81], and a trend toward greater ostracism intention than the moderate target, $t(266) = 1.69$, $p = .092$, $d = 0.34$, 95% CI = [-0.07, 0.74].

As in Study 3, liking toward the target was a major predictor of ostracism, $F(2, 265) = 86.12$, $p < .001$, $\eta_p^2 = .25$. After controlling for liking, the effect of condition became non-significant, $F(2, 265) = 0.54$, $p = .582$, $\eta_p^2 = .004$, and the gender interaction became marginal, $F(2, 265) = 2.52$, $p = .083$, $\eta_p^2 = .02$.

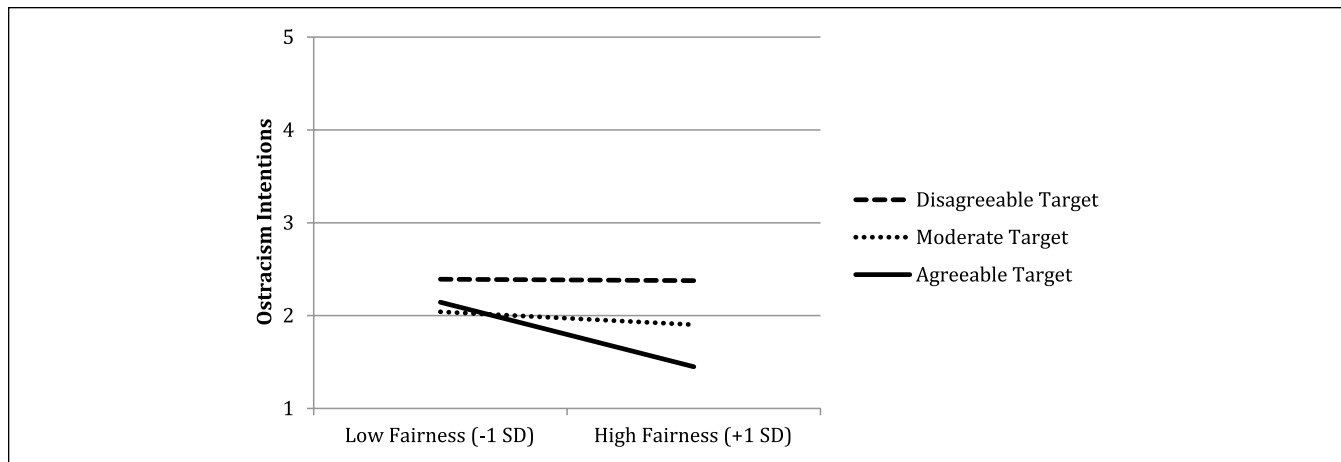
Fairness concerns. Replicating Study 3, participants' endorsement of fairness as a moral foundation moderated the effect of target agreeableness, $F(2, 267) = 5.69$, $p = .004$, $\Delta R^2 = .04$ (see Table 5). Higher fairness concerns predicted reduced intentions to ostracize the agreeable target, $b = -0.48$,

Table 5. Hierarchical Multiple Regression Analysis Predicting Ostracism Intentions From the Fairness Moral Foundation and Its Interaction With Target Agreeableness in Study 4.

Predictor	Model 1			Model 2		
	<i>b</i>	<i>t</i> -value	95% CI	<i>b</i>	<i>t</i> -value	95% CI
Step 1						
Fairness	−0.18**	−3.05	[−0.29, −0.06]	−0.01	−0.10	[−0.21, 0.19]
Dummy Code 1 (disagreeable vs. agreeable)	−0.57***	−5.61	[−0.78, −0.37]	−0.59***	−5.84	[−0.79, −0.39]
Dummy Code 2 (disagreeable vs. moderate)	−0.41***	−4.09	[−0.61, −0.21]	−0.41***	−4.15	[−0.61, −0.22]
Step 2						
Dummy Code 1 × Fairness				−0.47**	−3.15	[−0.77, −0.18]
Dummy Code 2 × Fairness				−0.09	−0.64	[−0.36, 0.18]
<i>R</i> ²	.13***	<i>F</i> = 13.99		.17***	<i>F</i> = 19.68	
<i>R</i> ² change				.04*	<i>F</i> = 5.69	

Note. Dummy Code 1 contrasts the agreeable and disagreeable conditions. Dummy Code 2 contrasts the moderate and disagreeable conditions. Betas are unstandardized, and fairness is mean centered. CI = confidence interval.

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

**Figure 2.** The relationship between endorsement of the fairness foundation and intentions to ostracize an agreeable, disagreeable, or control target in Study 4 ($n = 273$).

$t(267) = -4.45, p < .001$, but not the moderate target, $b = -0.10, t(267) = -1.09, p = .273$, or disagreeable target, $b = -0.01, t(267) < 0.10, p = .923$ (see Figure 2).

Examining the simple effects between all three conditions at different levels of fairness concerns, we found that participants who are low in fairness concerns (at 1 standard deviation below the mean) reported greater intentions to ostracize the disagreeable target relative to the moderate target, $b = -0.35, t(267) = -2.47, p = .014$, and (marginally) relative to the agreeable target, $b = -0.25, t(267) = -1.71, p = .089$. Responses to the moderate and agreeable targets did not differ, $b = 0.10, t(267) = -0.73, p = .467$. In contrast, those who are high in fairness concerns (at 1 standard deviation above the mean) reported significantly greater intentions to ostracize the disagreeable target than the agreeable target, $b = 0.93, t(267) = 6.20, p < .001$, and the moderate target, $b = -0.47,$

$t(267) = -3.47, p = .001$. They also reported significantly greater intentions to ostracize the moderate target than the agreeable target, $b = -0.45, t(267) = -3.14, p = .002$.

This interaction became marginally significant after controlling for liking, $F(2, 266) = 2.66, p = .072, \Delta R^2 = .01$.

Mediation analyses. We tested trust as a mediator of the effect of the descriptions of the target's personality on intentions to ostracize. Following Hayes and Preacher (2014), we created two dummy codes, with the disagreeable condition as the reference group. This allows us to test interpersonal trust as a mediator of both the agreeable–disagreeable contrast, and also the moderate–disagreeable contrast.

Both the agreeable and moderate targets were trusted more than the disagreeable target, $b = 0.82, t(270) = 10.33, p < .001$, and $b = 0.59, t(270) = 7.47, p < .001$, respectively.

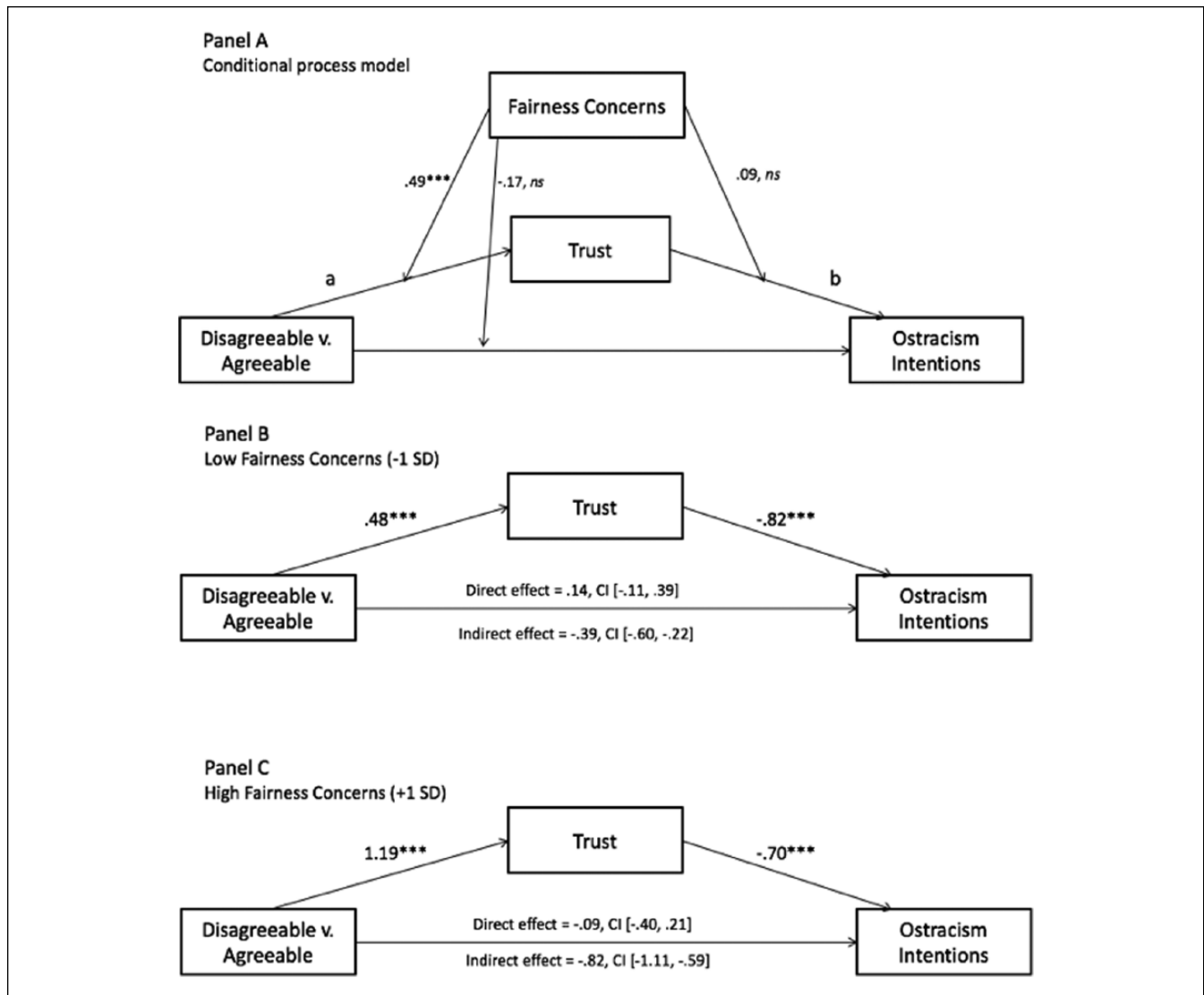


Figure 3. Indirect effect of target agreeableness on ostracism intentions through interpersonal trust, conditioned on fairness concerns (controls: disagreeable vs. moderate contrast, and the disagreeable vs. moderate contrast by fairness interaction term; $n = 273$).

Note. CI = confidence interval.

*** $p < .001$.

Trust, in turn, predicted less intention to ostracize, $b = -0.77$, $t(270) = -12.03$, $p < .001$. The indirect effect of trust was significant for both the agreeable–disagreeable contrast, *indirect effect* = -0.63 , bias-corrected 95% CI = $[-0.79, -0.49]$, and the moderate–disagreeable contrast, *indirect effect* = -0.45 , bias-corrected 95% CI = $[-0.59, -0.33]$. Critically, these indirect pathways through trust remained significant after controlling for liking, *indirect effect* = -0.16 , bias-corrected 95% CI = $[-0.27, -0.08]$, and *indirect effect* = -0.15 , 95% CI = $[-0.24, -0.07]$, respectively.

We conducted a conditional process analysis (Hayes, 2013, Process Model 59) to test (a) whether the indirect effect of the agreeable–disagreeable contrast is greater for participants who are high in fairness concerns, and if so (b)

whether it is due to a greater distrust of disagreeable participants (moderation of the *a* path), a greater effect of distrust on ostracism intentions (moderation of the *b* path), or both (see Panel A in Figure 3). For ease of exposition, our model focuses on the effects of the strongest contrast (agreeable vs. disagreeable), while controlling for the moderate–disagreeable contrast and its interaction with fairness concerns.

The analysis reveals that the indirect effect through trust is in fact contingent on participant's fairness concerns. The indirect pathway is significant even for those low in fairness concerns (-1 SD), indirect effect = -0.39 , 95% CI = $[-0.60, -0.22]$; however, it is stronger for those high in fairness concerns ($+1$ SD), indirect effect = -0.82 , 95% CI = $[-1.11, -0.59]$ (see Panels B and C in Figure 3, respectively).

An examination of the specific pathways reveals why this is the case. The effect of the disagreeable–agreeable contrast on interpersonal trust is stronger for participants who are highly concerned with fairness, and relatively weaker for those who are less concerned with fairness, interaction $b = 0.49$, $t(267) = 4.32$, $p < .001$. However, trust strongly predicted ostracism intentions regardless of people’s moral fairness concerns, interaction $b = 0.09$, $t(265) = 1.03$, $p = .302$.

We conducted the same analysis, but controlling for liking and its interaction with fairness concerns. In this model, the interaction between fairness concerns and the disagreeable–agreeable contrast in predicting trust remained significant, $b = 0.24$, $t(265) = 2.32$, $p = .021$. The indirect effect of trust also remained significant controlling for liking at high levels of fairness concerns, *indirect effect* = $-.24$, 95% CI = $[-0.45, -0.10]$, but was not significant at low levels of fairness concerns, *indirect effect* = $-.09$, 95% CI = $[-0.22, 0.04]$.

Together this analysis suggests that reduced interpersonal trust is a mechanism by which disagreeableness leads to ostracism.

Discussion

Both males and females showed intentions to ostracize a target described as low in agreeableness compared with a target described as high in agreeableness. However, relative to the moderately agreeable target, only females showed an additional *decrease* in intentions to ostracize an agreeable target. This finding suggests that females, who are generally more oriented toward communal goals (Diekmann, Clark, Johnston, Brown, & Steinberg, 2011), do not merely use social ostracism as a stick to punish the disagreeable, but may also use social inclusion as a carrot to reward the agreeable.

The analysis of the fairness moral foundation showed that the moderating effect discovered in Study 3 is reliable; fairness concerns predict intentions to ostracize agreeable targets, but not disagreeable or moderate targets, implying that the effect of target personality on ostracism intentions is motivated in part by a preference for non-exploitative exchange partners. The moderated mediation analysis provided direct support for this conclusion; the disagreeable target was trusted less and subsequently ostracized more, especially in participants high in fairness concerns, even after controlling for liking.

Interestingly, fairness concerns moderated how much disagreeable targets were distrusted, but not how much this distrust predicted ostracism. This suggests that people high in fairness concerns are not only morally concerned with fairness violations that have occurred (as articulated by moral foundation theory, Graham et al., 2009) but they may also be especially vigilant in detecting possible future violations of fairness, as seen in their greater feelings of distrust toward disagreeable participants.

These three experiments offer varying levels of support for an effect of agreeableness above and beyond participant’s

liking of the target, with Study 3 finding evidence for such an effect, but not so in Studies 2 and 4. This is likely due to some conceptual overlap between agreeableness and likeability. Early models of personality labeled the agreeableness dimension as “friendly compliance,” or “possessing a pleasant disposition,” with one even applying the label “likeability” as a name for the dimension (Hogan, 1983; for a review, see Graziano & Eisenberg, 1997). In line with this conceptual similarity is the empirical similarity observed in these three studies. Perceptions of agreeableness shared a great amount of variance with likeability ratings (r s from .55 to .64). Considering the conceptual and empirical closeness, we were impressed there is any support for an effect of agreeableness beyond liking.

Study 4, however, provides evidence for an additional pathway between agreeableness and ostracism: interpersonal trust. This pathway was robust even after controlling for liking, suggesting that disagreeableness leads to ostracism, in large part because disagreeable people are disliked but also because they are distrusted. Importantly, the moderating effect of fairness concerns was not significant after controlling for liking.

It is also worth noting that even though disliking the target is a somewhat unsurprising cause of ostracism, this link has yet to be documented in the literature. Existing accounts suggest that burdensome members are ostracized, as are those who threaten group identity (Kurzban & Leary, 2001). However, the current findings suggest people may not be so principled in choosing whom to ostracize. A non-burdensome member (e.g., one carrying his or her weight) may still be ostracized simply for being disagreeable. This is interesting because ostracizing others can be costly (Legate et al., 2013). One might expect people to reserve ostracism for more serious social violations than merely being disliked. However, the strong effects of disliking on ostracism in the current studies suggest that disliking may be sufficient.

Studies 2 to 4 show that target disagreeableness can lead to greater intentions to ostracize. We now turn our attention to the reverse causal pathway. Does ostracism lead to a temporary disagreeable state? Studies 5 and 6 tested this hypothesis by randomly assigning participants to be ostracized or included before responding to a state measure of agreeableness.

Study 5

Method

Participants and design. Participants in public areas on the Purdue University campus were invited to participate in a brief study of mental visualization conducted on an iPad. Individuals ($N = 86$) were approached by a female experimenter and invited to participate in a brief study of mental visualization. Of these, 56 agreed to participate (21 males; $M_{\text{Age}} = 21.30$, $SD = 5.14$, participation rate = 66.28%). Four

participants were excluded from the analysis because they were familiar with Cyberball, leaving a total sample of 52. Participants were randomly assigned to be ostracized or included.

Procedure. Participants were told the purpose of the experiment was to examine mental visualization. They were asked to play an online ball-tossing game called Cyberball, played with two agents that participants are lead to believe are real players. In reality, they are programmed to differentially include the participant based on condition (Williams, Cheung, & Choi, 2000). Participants in the inclusion condition ($n = 26$) received a fair number of throws (33%) throughout the experience. Participants in the ostracism condition ($n = 26$) received two throws at the beginning of the game, and none thereafter. The game lasted 21 throws and was played for approximately 2 min. Following the game, participants completed a short measure unrelated to the current research.

Measures. After playing Cyberball, participants completed a modified version of the nine agreeableness questions from the Big Five Inventory (John et al., 1991) intended to assess the level of agreeableness at that moment, rather than dispositional agreeableness. Participants read, "The following statements refer to how you would feel and act around the people that you generally spend time with if they were present at this moment. Please rate your agreement with the following statements based on how you are feeling right now." They then responded to the nine items used to assess agreeableness in Studies 1 and 2, only this time in reference to how they currently felt (e.g., "Right now I feel like I would be helpful and unselfish with others"; $\alpha = .91$).

As a manipulation check, all participants completed two items rating the extent to which they were (a) ignored and (b) excluded during the game, which were averaged together ($r = .96$). They also estimated the percentage of throws received.

Results and Discussion

Manipulation checks. Ostracized participants reported being more ignored and excluded ($M = 4.46$, $SD = 0.86$), than included participants ($M = 2.04$, $SD = 0.99$), $t(50) = -9.38$, $p < .001$, $d = 2.61$, 95% CI = [1.86, 3.35]. They also reported receiving a smaller percentage of ball tosses ($M = 10.04$, $SD = 4.90$) than included participants ($M = 34.90$, $SD = 11.21$), $t(43) = 9.98$, $p = .007$, $d = 2.87$, 95% CI = [2.02, 3.71].

Agreeableness. Ostracized participants reported lower state agreeableness ($M = 3.29$, $SD = 0.85$) than included participants ($M = 3.92$, $SD = 0.67$), $t(50) = 2.98$, $p = .004$, $d = 0.82$, 95% CI = [0.25, 1.38].

Study 5 provides evidence that ostracism induces disagreeableness. To identify the mechanism underlying this effect, in Study 6, we replicated these procedures while also measuring three variables known to be affected by ostracism:

needs satisfaction, positive affect, and negative affect (both anger and sadness). Although we did not have a clear hypothesis regarding needs satisfaction or positive affect, we did expect that anger would lead to more disagreeableness than sadness in a pattern similar to Chow et al. (2008).

Study 6

Method

Participants and design. Participants were recruited from public areas on campus. Individuals ($N = 116$) were invited to participate, and 92 agreed (39 males; $M_{Age} = 20.79$, $SD = 2.13$, participation rate = 79.31%). Seventeen participants were excluded from analysis because they were familiar with Cyberball, leaving a total sample of 75.

The procedure was identical to Study 5, except before reporting agreeableness, participants reported needs satisfaction, positive affect, and negative affect. Separate items were used to assess anger and sadness, allowing us to test the differential mediating role played by each emotion.

Measures. Needs satisfaction was measured with eight items assessing the four basic needs: belonging (e.g., "During the game I felt rejected," reverse scored), self-esteem (e.g., "During the game I felt good about myself"), control (e.g., "During the game I felt powerful"), and meaningful existence (e.g., "During the game I felt invisible," reverse scored). These items were averaged, forming a composite needs satisfaction index ($\alpha = .88$). Positive affect was measured with the three items ("good," "friendly," and "relaxed"; $\alpha = .74$). Negative affect was measured with five items ("bad," "unfriendly," "angry," "sad," and "tense"; $\alpha = .90$). These were adapted from the scale provided by Williams (2009), with added items, "tense" and "relaxed." Participants responded on a scale from 1 (*not at all*) to 5 (*completely*).

Results

Manipulation checks. Ostracized participants correctly reported being more ignored and excluded ($M = 4.03$, $SD = 1.02$) than included participants ($M = 1.78$, $SD = 0.88$), $t(73) = -10.19$, $p < .001$, $d = 2.36$, 95% CI = [1.76, 2.95]. They also reported receiving fewer ball tosses ($M = 13.37$, $SD = 10.57$) than included participants ($M = 34.51$, $SD = 10.61$), $t(68) = 8.75$, $p < .001$, $d = 2.09$, 95% CI = [1.50, 2.67].

Ostracism outcomes. Relative to included participants, those who were ostracized reported lower needs satisfaction, $t(73) = 6.98$, $p < .001$, $d = 1.61$, 95% CI = [1.08, 2.13], lower positive affect, $t(73) = 3.32$, $p = .001$, $d = 0.76$, 95% CI = [0.29, 1.23], higher negative affect, $t(73) = -5.45$, $p < .001$, $d = -1.26$, 95% CI = [-1.75, -0.76], and marginally lower agreeableness, $t(73) = 1.87$, $p = .07$, $d = 0.43$, 95% CI = [-0.03, 0.89] (see Table 6).

Table 6. Means and Standard Deviations for Ostracism Outcomes by Condition in Study 6 ($n = 75$).

Dependent variable	Condition	
	Included	Ostracized
Need satisfaction	3.53 [3.31, 3.76]	2.34 [2.08, 2.60]
Anger	1.59 [1.27, 1.92]	2.71 [2.32, 3.10]
Sadness	1.70 [1.36, 2.04]	2.87 [2.46, 3.27]
Agreeableness	3.82 [3.61, 4.04]	3.55 [3.33, 3.77]

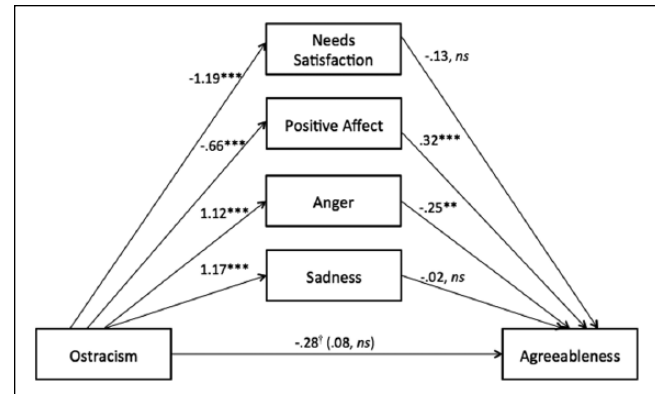
Note. Responses for all items are on a 5-point scale. 95% confidence intervals appear in brackets.

Mediation analysis. We tested the various ostracism outcomes as mediators in two stages. First, we conducted a multiple mediation model testing the aggregated needs satisfaction, positive affect, and negative affect scales each as simultaneous mediators. Agreeableness was predicted by positive affect, $b = 0.29$, $t(70) = 3.14$, $p = .002$, and negative affect, $b = -0.29$, $t(70) = -3.46$, $p < .001$, but not needs satisfaction, $b = -0.10$, $t(70) = -0.75$, $p = .406$. A bias-corrected bootstrap 95% CI for the total indirect effect ($-.40$) based on 5000 bootstrap samples did not include zero, *indirect effect* = $-.40$ 95% CI = $[-0.67, -0.16]$.

Next, to examine anger and sadness as particular affective states of interest, we conducted a follow-up mediation in which these two items were entered individually, alongside the aggregate needs satisfaction and positive affect scales (see Figure 4). Agreeableness was predicted by positive affect, $b = 0.32$, $t(69) = 3.53$, $p < .001$, and anger, $b = -0.25$, $t(69) = -3.16$, $p = .002$, but not needs satisfaction, $b = -0.13$, $t(69) = -1.11$, $p = .270$, or sadness, $b = -0.02$, $t(69) = -0.258$, $p = .797$. The indirect pathways were significant for both positive affect, *indirect effect* = $-.21$, 95% CI = $[-0.44, -0.07]$, and anger, *indirect effect* = $-.28$ $[-.62, -.07]$ (see Figure 4). This was also the case in a final model that also included the remaining negative affect items. The only significant indirect pathways were through anger, *indirect effect* = $-.33$ 95% CI = $[-0.74, -0.08]$, and positive affect, *indirect effect* = $-.23$, 95% CI = $[-0.48, -0.08]$.

Discussion

These findings show that the negative affect resulting from ostracism has the downstream consequence of inducing a state of disagreeableness. More specifically, however, it appears that feelings of anger, rather than sadness, account for the disagreeable state caused by ostracism. This conclusion aligns with research showing the powerful effect of anger in guiding people toward antisocial behaviors following ostracism (Chow et al., 2008). The effect was also

**Figure 4.** Multiple mediation model testing the indirect effects of ostracism on agreeableness (Study 6; $n = 75$).

Note. Coefficients are unstandardized. The coefficient in parentheses represents the effect when the mediators are added to the model. The adjacent coefficient represents the direct effect of ostracism on agreeableness.

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

explained by ostracism's role in inducing positive affect. We speculate the positive affect may be especially useful in preparing individuals for cooperative interactions. We note, however, that these analyses, while consistent with a causal mediating effect of sadness and positive affect, are limited by their correlational nature.²

General Discussion

Based on these findings, we conclude that ostracism is associated with disagreeableness, disagreeableness elicits ostracism from others, and ostracism provokes a state of disagreeableness that is mediated by the affective consequences of ostracism. In combination, this research suggests that the very characteristic that can cause someone to be ostracized is temporarily amplified by the ostracism event. This pattern is reminiscent of behavioral confirmation processes, which are prominently featured in social psychology (Snyder, 1984). Just as perceived intelligence can lead to actual intelligence (Rosenthal & Jacobson, 1968), and perceived liking can lead to actual liking (Snyder, Tanke, & Berscheid, 1977), so too can ostracism based on disagreeableness lead to a state of actual disagreeableness.

Implications for Theories of Personality

This research contributes to a growing literature suggesting that personality can and does change (Roberts, Walton, & Viechtbauer, 2006). In the broadest sense, mean-level change in personality traits across the life span is well documented (Roberts & Mroczek, 2008). Recent research also provides evidence that career path and both positive and negative life events also play a significant role in personality change over time (Lüdtke, Roberts, Trautwein, & Nagy, 2011). We argue

that chronic ostracism represents a powerfully negative life event, and has the potential to effect significant, lasting personality change. This is partially supported by our first study, which provided evidence that trait agreeableness is correlated negatively with self-report chronic ostracism. The correlational nature of this finding, however, limits our ability to make claims about the directional nature of chronic ostracism's effect on trait-level agreeableness.

We do, however, also provide evidence that a single ostracism experience decreases state agreeableness. Fleeson (2001) has argued that personality should be conceptualized as "density distributions" of traits. Across social interactions and situational contexts, each individual will express multiple levels of a single trait, rather than one fixed level. This personal distribution of traits expressions may be a better depiction of an individual's personality than a mean-level trait score, because it includes information about variability, or reactivity to situational factors. From this perspective, evidence that a single instance of ostracism leads to the expression of lower levels of the trait agreeableness carries clear implications. The frequency of ostracism experiences has the potential to shift the mean level of any individual's distribution of expressed agreeableness, or the potential to increase within-person variance on agreeableness, resulting in lower trait-level stability. This has meaningful implications for the long-term impact of ostracism.

Effortful Control

As we previously discussed, ostracized individuals display a variety of responsive behaviors: some prosocial, some antisocial (Smart Richman & Leary, 2009). We proposed that decreased agreeableness associated with experienced ostracism should lead primarily to antisocial behavioral responses to rejection. Nonetheless, prosocial efforts following social exclusion are well documented. One potential explanation is that prosocial responses to ostracism represent an effortful override of an initial antisocial impulse. Such an override likely represents the expression of effortful control, a temperament system associated with self-regulation (Rothbart, Derryberry, & Posner, 1994). Effortful control, the ability to replace a dominant response with a non-dominant, but more socially desirable response, may be largely predictive of decisions to engage in prosocial rather than aggressive or antisocial behavior in response to ostracism (Kochanska, Murray, & Coy, 1997). Effortful control has been linked to the regulation of aggression in children (Murray & Kochanska, 2002) and may play a role in similar regulation in adults. Indeed, Jensen-Campbell, Rosselli, et al. (2002) suggest that effortful control is related to trait agreeableness and conscientiousness in adults. It seems agreeableness may be, in part, necessary for the override of antisocial behavior in response to ostracism (Graziano & Habashi, 2010). That ostracism lowers agreeableness, as evidenced here, suggests that exposure to chronic ostracism may eventually push

agreeableness levels below a threshold necessary for the effortful replacement of antisocial behavior with prosocial behavior. For now, these ideas are speculative; a more thorough examination of the role of effortful control is an important direction for future research.

Such a process would be consistent with existing reconciliations of pro versus anti social behavior following ostracism. For example, aggressive responses are thought to be more likely when control and meaningful existence are threatened (the power-provocation needs; Williams, 2009). Threats to these needs may immediately incite aggressive impulses, which are more likely to be realized in part due to compromised agreeableness and associated effortful control. Similarly, aggressive responses are thought to be more likely when the relationship is not valued and when the rejection is seen as especially unfair (Smart Richman & Leary, 2009). Ordinarily, people may be able to behave appropriately even toward interaction partners in less important relationships (i.e., a casual acquaintance vs. a friend); however, in the throws of an ostracism-induced disagreeable state, they may lack the effortful control required to not aggress against sources of ostracism in less important relationships.

Limitations

To test whether disagreeableness elicits ostracism from others, we asked sources to respond to vignettes, rather than observe their actual behavior. It is possible that the effect of disagreeableness on ostracism is smaller in face-to-face situations where the target's pain is visible to the source. Indeed, research shows that ostracism can be an unpleasant event, even for the perpetrator (Legate et al., 2013). However, much of the literature showing that sources experience unpleasant reactions when ostracizing comes from studies that ask participants to ostracize when the participants have no justification for doing so. Thus, it is likely that when sources are motivated to ostracize, they may be less attentive to the pain they inflict.

Second, in testing whether ostracism provokes disagreeableness, this research assessed state, rather than trait agreeableness. These findings do not directly test the effects of ostracism on global agreeableness self-evaluations or the effects of ostracism over time. The decision to measure state agreeableness was guided by the relatively minimal nature of Cyberball. It is unlikely that such a brief ostracism episode would alter people's global perceptions of how agreeable they are. Instead, this research demonstrates that a single ostracism event leads at least to a momentary state of disagreeableness. One might reasonably extrapolate that just as one ostracism experience leads to a state of disagreeableness, so too should many ostracism experiences lead to many states of disagreeableness. Likewise, chronic ostracism can be expected to lead to chronic (trait) disagreeableness. We note that we did not account for people's baseline agreeableness in Studies 4 through 7. An intriguing possibility for

future research is that people who are disagreeable are less resilient in the face of an ostracism episode, and thus more prone to the temporary disagreeable states shown to be provoked by ostracism.

In addition, it is possible that the effects of Cyberball on state agreeableness were due to *increased agreeableness* in participants who were included. This is possible, but given the strong social norm to include others, participants generally expect to be included in Cyberball. Accordingly, studies typically find little difference between inclusion and control conditions (e.g., Riva et al., 2014).

Conclusion

This research demonstrates that agreeableness can play a role both as a cause and a consequence of ostracism. In addition to other documented negative outcomes associated with disagreeableness, it appears the disagreeable are also vulnerable to ostracism. In fact, given that disagreeableness leads to ostracism and also negative life outcomes (Ozer & Benet-Martínez, 2006), one might wonder if ostracism and subsequent need threat may help account for the reduced longevity, job attainment, and psychological health associated with disagreeableness.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. An ANCOVA controlling for the manipulation check verified that this interaction was not due to gender differences in detection of agreeableness, $F(2, 265) = 3.95, p = .020, \eta_p^2 = .03$.
2. In addition to Studies 5 and 6, we also conducted a lab study ($n = 52$) in which ostracism marginally decreased state agreeableness, $t(50) = 1.92, p = .061, d = 0.53$, confidence interval (CI) = $[-0.03, 1.08]$. A meta-analysis of the three studies (Cumming, 2012) indicated that ostracism significantly decreases agreeableness, $d = 0.57, 95\% \text{ CI} = [0.27, 0.87]$.

Supplemental Material

The online supplemental material is available at <http://pspb.sagepub.com/supplemental>.

References

- Berkowitz, L., & Harmon-Jones, E. (2004). Toward an understanding of the determinants of anger. *Emotion, 4*, 107-130.
- Carter-Sowell, A. R. (2010). *Salting a wound, building a callous, or throwing in the towel? The measurement and effects of chronic ostracism experiences* (Unpublished doctoral dissertation). Purdue University, West Lafayette, IN.
- Chow, R. M., Tiedens, L. Z., & Govan, C. L. (2008). Excluded emotions: The role of anger in antisocial responses to ostracism. *Journal of Experimental Social Psychology, 44*, 896-903. doi:10.1016/j.jesp.2007.09.004
- Cumming, G. (2012). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. New York, NY: Routledge.
- Diekmann, A. B., Clark, E. K., Johnston, A. M., Brown, E. R., & Steinberg, M. (2011). Malleability in communal goals and beliefs influences attraction to stem careers: Evidence for a goal congruity perspective. *Journal of Personality and Social Psychology, 101*, 902-918.
- Feinberg, M., Willer, R., & Schultz, M. (2014). Gossip and ostracism promote cooperation in groups. *Psychological Science, 25*, 656-664. doi:10.1177/0956797613510184
- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*, 1011-1027.
- Gómez, Á., Morales, J., Hart, S., Vázquez, A., & Swann, W. R. (2011). Rejected and excluded forevermore, but even more devoted: Irrevocable ostracism intensifies loyalty to the group among identity-fused persons. *Personality and Social Psychology Bulletin, 37*, 1574-1586. doi:10.1177/0146167211424580
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different moral foundations. *Journal of Personality and Social Psychology, 96*, 1029-1046.
- Graziano, W. G., & Eisenberg, N. (1997). Agreeableness: A dimension of personality. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 795-824). San Diego, CA: Academic Press.
- Graziano, W. G., & Habashi, M. M. (2010). Motivated processes underlying both prejudice and helping. *Personality and Social Psychology Review, 14*, 313-331.
- Graziano, W. G., Habashi, M. M., Sheese, B. E., & Tobin, R. M. (2007). Agreeableness, empathy, and helping: A person \times situation perspective. *Journal of Personality and Social Psychology, 93*, 583-599. doi:10.1037/0022-3514.93.4.583
- Graziano, W. G., & Tobin, R. M. (2009). Agreeableness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 46-61). New York: Guilford
- Gruter, M., & Masters, R. (1986). Ostracism as a social and biological phenomenon: An introduction. *Ethology and Sociobiology, 7*, 227-236.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: The Guilford Press.
- Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology, 67*, 451-470. doi:10.1111/bmsp.12028
- Hogan, R. T. (1983). A socioanalytic theory of personality. In M. Page (Ed.), *Nebraska symposium on motivation: Personality-current theory and research* (pp. 58-89). Lincoln: University of Nebraska Press.
- Jensen-Campbell, L. A., Adams, R., Perry, D. G., Workman, K. A., Furdella, J. Q., & Egan, S. K. (2002). Agreeableness, extraversion, and peer relations in early adolescence: Winning friends and deflecting aggression. *Journal of Research in Personality, 36*, 224-251. doi:10.1006/jrpe.2002.2348

- Jensen-Campbell, L. A., Rosselli, M., Workman, K. A., Santisi, M., Rios, J. D., & Bojan, D. (2002). Agreeableness, conscientiousness, and effortful control processes. *Journal of Research in Personality, 36*, 476-489.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102-138). New York: Guilford Press.
- Johnson-George, C., & Swap, W. C. (1982). Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology, 43*, 1306-1317. doi:10.1037/0022-3514.43.6.1306
- Kagel, J., & McGee, P. (2014). Personality and cooperation in finitely repeated prisoner's dilemma games. *Economics Letters, 124*, 274-277.
- Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development, 68*, 263-277.
- Kurzban, R., & Leary, M. R. (2001). Evolutionary origins of stigmatization: The functions of social exclusion. *Psychological Bulletin, 127*, 187-208.
- Legate, N., DeHaan, C. R., Weinstein, N., & Ryan, R. M. (2013). Hurting you hurts me too: The psychological costs of complying with ostracism. *Psychological Science, 24*, 583-588.
- Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion, 14*, 473-493. doi:10.1080/026999300402763
- Lüdtke, O., Roberts, B. W., Trautwein, U., & Nagy, G. (2011). A random walk down university avenue: Life paths, life events, and personality trait change at the transition to university life. *Journal of Personality and Social Psychology, 101*, 620-637.
- Murray, K., & Kochanska, G. (2002). Effortful control: Factor structure and relation to externalizing and internalizing behaviors. *Journal of Abnormal Child Psychology, 30*, 503-514.
- Ozer, D. J., & Benet-Martinez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology, 57*, 401-421. doi:10.1146/annurev.psych.57.102904.190127
- Riva, P., Williams, K. D., Torstrick, A. M., & Montali, L. (2014). Orders to shoot (a camera): Effects of ostracism on obedience. *The Journal of Social Psychology, 154*, 208-216. doi:10.1080/00224545.2014.883354
- Roberts, B. W., & Mroczek, D. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*, 31-35.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*, 1-25. doi:10.1037/0033-2909.132.1.1
- Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectations and student intellectual development*. New York, NY: Holt, Rinehart and Winston.
- Rothbart, M. K., Derryberry, D., & Posner, M. I. (1994). A psychological approach to the development of temperament. In J. E. Bates & T. D. Wachs (Eds.), *Temperament: Individual differences at the interface of biology and behavior* (pp. 83-116). Washington, DC: American Psychological Association.
- Smart Richman, L. S., & Leary, M. R. (2009). Reactions to discrimination, stigmatization, ostracism, and other forms of interpersonal rejection: A dynamic, multi-motive model. *Psychological Review, 116*, 365-383.
- Snyder, M. (1984). When belief creates reality. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 18, pp. 247-305). Orlando, FL: Academic Press.
- Snyder, M., Tanke, E. D., & Berscheid, E. (1977). Social perception and interpersonal behavior: On the self-fulfilling nature of social stereotypes. *Journal of Personality and Social Psychology, 35*, 656-666. doi:10.1037/0022-3514.35.9.656
- Twenge, J. M., Baumeister, R. F., DeWall, C., Ciarocco, N. J., & Bartels, J. (2007). Social exclusion decreases prosocial behavior. *Journal of Personality and Social Psychology, 92*, 56-66. doi:10.1037/0022-3514.92.1.56
- Warburton, W. A., Williams, K. D., & Cairns, D. R. (2006). When ostracism leads to aggression: The moderating effects of control deprivation. *Journal of Experimental Social Psychology, 42*, 213-220.
- Wesselman, E. D., Butler, F. A., Williams, K. D., & Pickett, C. L. (2010). Adding injury to insult: Unexpected rejection leads to more aggressive responses. *Aggressive Behavior, 36*, 232-237. doi: 10.1002/ab.20347
- Wesselmann, E. D., Wirth, J. H., Pryor, J. B., Reeder, G. D., & Williams, K. D. (2013). When do we ostracize? *Social Psychological & Personality Science, 4*, 108-115.
- Williams, K. D. (1997). Social ostracism. In R. Kowalski (Ed.), *Aversive interpersonal behaviors* (pp. 133-170). New York, NY: Plenum.
- Williams, K. D. (2009). Ostracism: Effects of being excluded and ignored. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 41, pp. 275-314). New York, NY: Academic Press.
- Williams, K. D., Cheung, C., & Choi, W. (2000). Cyberostracism: Effects of being ignored over the internet. *Journal of Personality and Social Psychology, 79*, 748-762.
- Williams, K. D., & Wesselmann, E. D. (2011). The link between ostracism and aggression. In J. P. Forgas, A. W. Kruglanski, & K. D. Williams (Eds.), *The psychology of social conflict and aggression* (pp. 37-51). New York, NY: Psychology Press.