Comparative Humility:

How The Perceived Necessity of Moral Qualities Explains The Individuation Effect

Submitted on: May 13, 2019

Abstract

Most people think they are above average, a statistically impossible result. That said, people compare themselves less favorably to a specific individual than they do to the population from which that individual was drawn. We seek to explain this *individuation effect*, a previously uncovered phenomenon that has been both influential and mysterious. New data (and a reanalysis of published results) show this individuation effect emerges to the extent a trait is naturally viewed or experimentally framed as moral. People explicitly indicate a desire to give others the benefit of the doubt until proven wrong on moral dimensions. But because people apply this expressed strategy only in judgments against specific people, not people in general, the individuation effect emerges. People adopt this "benefit of the doubt" strategy because they view moral qualities as necessary for others to possess; their absence is a dealbreaker for social investment. In this way, people avoid preemptively dismissing potential social investments (i.e., specific people instead of people in general) for failing to bring what the self can to social relationships, thereby remaining open-minded about the potential worth of new social ties. This illustrates how people can maintain their sense of high standing in a population without seeing specific social ties as unworthy of investment.

KEYWORDS: social comparison, better-than-average effect, morality, humility, self-enhancement

Comparative Humility:

How The Perceived Necessity of Moral Qualities Explains The Individuation Effect

Defying logic and statistics, too many people claim they are better than the average person (Alicke & Govorun, 2005; Dunning, 2005; Dunning, Heath, & Suls, 2004). Historically, explanations for this better-than-average effect, or illusory superiority, have focused on why the self is evaluated so positively. People may distort impressions to flatter the self (Brown & Han, 2012; Guenther & Alicke, 2010), simply be blind to their own shortcomings (Ehrlinger, Johnson, Banner, Dunning, & Kruger, 2008; Kruger & Dunning, 1999), enjoy flattering self-characterizations while accountability is distant or unlikely (Sweeny & Krizan, 2013), redefine traits (Critcher, Helzer, & Dunning, 2011; Dunning, Meyerowitz, & Holzberg, 1989) or their performance potential (Klein & Goethals, 2002) to permit the self to outshine others, or simply justify their superiority due to greater confidence in their positive self-knowledge than in similar knowledge of others (Hilbert, 2012; Moore & Healy, 2008).

Other explanations have focused on what differentiates the self-perceptions of healthy, well adjusted individuals—those known to show especially strong positive illusions about the self (Taylor & Brown, 1988). Such individuals are faster to retrieve positive (vs. negative) autobiographical memories (see Blaney, 1986), possess positively distorted recall of their past performance (Crary, 1966), and show an inhibition of contradictory episodic memories when considering their own standing on positive traits (Hitchcock, Rees, & Dalgleish, 2017). These explanations are not mutually exclusive, but their sheer number shows the attention the self has received in this research tradition (Alicke & Sedikides, 2011).

These explanations, however, all share a characteristic: They explain self-illusory beliefs by focusing primarily on the impressions people hold about the self, with a concern for the

comparison other being something of an afterthought (Kruger, 1999). In considering the comparison other, the literature has mostly been silent. One exception, however, stands out. Alicke et al. (1995) found that undergraduate participants compared themselves less favorably to a single randomly-selected student than to students in general, even though one could argue that both judgments should be equivalent. This pattern is the *individuation effect*. Across multiple studies, this (minimally) individuated student was instantiated in several ways—e.g., as the stranger "sitting next to you," a person displayed on a TV screen, or as a person whose interview transcript one read. "The degree of reduction due to individuation was approximately the same across these various conditions" (p. 823; Alicke et al., 1995). That Alicke and colleagues' paper has over one thousand Google Scholar citations speaks to how influential these findings has been in the literature on illusory superiority, but the psychology behind them has remained largely mysterious.

In this manuscript, we turn our attention to this individuation effect and ask why, when, and how it might arise. We propose that this individuation effects reflects an asymmetric strategy people follow when considering comparison others—a strategy that applies differently when considering a specific person versus people in general. We begin by noting that social interactions occur with specific individuals, not with entire populations. In those interactions, certain traits provide the foundation upon which social life thrives (Haidt & Kesebir, 2010).

Herein, however, lies a potential problem. If the self sees itself as much better than others, this could discourage the formation of successful social bonds with other people. After all, within relationships, feeling superior to another individual is associated with disliking that person more, and also being disliked more, especially over the long-term (Colvin, Block, & Funder, 1995; Kwan, John, Robins, & Kuang, 2008; Paulhus, 1998; Rentzsch & Schröder-Abé,

2015). Further, if one approaches potential social relationship partners with the idea that the self dominates the other on traits deemed essential to social relationships, the self would not be particularly motivated to test the waters of this potential partnership. Thus, we propose that the self tempers its relative self-aggrandizement when comparing itself to specific individuals, especially on those traits where seeing the self as much better than a specific other could prove maladaptive and doom a possibly fruitful social investment before it begins. For traits of this type, we suggest people give others the benefit of the doubt until proven wrong.

If the self displays comparative humility when it thinks about its standing compared to another person (as opposed to people in general), what are these essential traits that may prompt this individuation effect? We argue that those qualities that are essential to social relationships are those that are moral in nature (Helzer & Critcher, 2018). Not only do moral qualities loom large in our assessments of the overall favorability of a person (Goodwin, Piazza, & Rozin, 2014), but they are highly predictive of whom we choose to engage in social exchange. Moral traits are the most important and meaningful within social interaction, if one believes both classic (Anderson, 1968) and more contemporary work (Wojciszke, Bazinska, & Jaworski, 1998) on which particular traits drive impressions of other people, as well as which traits feature more prominently in social compared to self assessments (Wojciszke, 1994). In building social bonds and feelings of solidarity, it is the traits associated with communion and morality that figure in most. When participants were asked to characterize in whom they would invest time and interpersonal resources, moral traits topped the list of necessary qualities; immoral qualities were dealbreakers (Hartley, Furr, Jayawickreme, Helzer, Vealsquez, & Fleeson, 2016). Traits associated with other characteristics, like competence and agency, are more relevant to the self and its private goals than they are for furthering interpersonal goals (Abele & Wojciszke, 2007).

By one argument, the social necessity of moral qualities might lead one to think that people should set especially high standards for identifying such qualities in others. And in fact, previous research on cynicism in social judgment would seem to support this perspective. For example, people seem quite prepared to attribute selfish motives to behaviors first identified as selfless (Critcher & Dunning, 2011; Newman, 2014). Such processes are in part responsible for maintaining a belief that self-serving attributions, intentions, and behaviors are more common than they actually are (Kruger & Gilovich, 1999; Ratner & Miller, 2001).

Although a suspicious vigilance for signs one is being duped or cheated may produce this skepticism about others' *past* behaviors (Cosmides, Tooby, Fiddick, & Bryant, 2005; Fein, McCloskey, & Tomlinson, 1997), we argue that a prospective cynicism about specific others would keep people from ever finding social partners with whom to engage. Consider the experience of a perspective homebuyer in one of the world's most expensive real estate markets. Week after week, open house after open house, the searcher returns to his apartment defeated. It would be easy for cynicism to set in and for the search to stop. But in that case, the search for a home would be doomed for sure. But if the searcher can maintain some prospective optimism—a belief that the next home on one's list may be the one—the search will continue; the pathway to the new home will remain open. Note that in this example, one need not maintain optimism about the housing market as a whole. Instead, one is well served by optimism about any specific house one is considering. There may be a useful time for a particularly skeptical eye before one places an offer on the home, but prospective optimism may be necessary to motivate this search to begin with.

Hints exist that this pattern is observable in people's approach to the morality of others. On the one hand, people possess cynical beliefs about the trustworthiness of their peers

(Fetchenhauer & Dunning, 2009, 2010). But they often go ahead and trust complete strangers—even though they expect a negative return for doing so (Dunning, Anderson, Schlösser, Ehlebracht, & Fetchenhauer, 2014; Fetchenhauer & Dunning, 2009, 2012). Crucial for the present analysis is that such prosocial behavior is not motivated by a dispassionate expected value calculation, but instead by a moral norm that one should respect the other person (Dunning, 2017; Dunning et al., 2014; Dunning, Fetchenhauer, & Schlösser, 2016). Applying this perspective to the present line of argumentation, people may feel they should not place themselves on a higher moral ground than others. Embedded in the Biblical adage *Let he who is without sin cast the first stone* is not merely an admonition against passing judgment on others. By invoking the social comparison in particular, it speaks to the hubris of placing the self on a higher moral pedestal than another.

Our central prediction is that the individuation effect—the reduction in self-enhancement that comes from comparisons against an individual compared to comparisons a group—will be driven by moral qualities. This is because the self will be reluctant to self-aggrandize against specific (even unknown) others on qualities it deems socially essential for people to possess. Moral qualities—given their key significance in social relationships—typify such essential qualities. By not dismissing individuals *ex ante* as morally inferior to the self (and thus an unsuitable match for social investment), the self should display comparative humility when comparing itself against individuals.

Preliminary Data

As an initial test of this idea, we returned to Alicke et al.'s (1995) Study 1. Participants rated themselves on 20 positive traits, comparing themselves to the average of the student population or to a same-sex stranger whom they saw but did not interact with. From descriptive

statistics reported in the article, we calculated for each trait the effect size (Cohen's *d*) of the individuation effect—i.e., the reduction in illusory superiority observed when comparing the self to a specific individual rather than to the population as a whole. We asked 126 members of an American university's on-line subject pool to rate each trait on a scale from 1 (*not related to moral character*) to 7 (*very reflective of moral character*). Consistent with our account, the strength of the individuation effect correlated very strongly with the trait's moral connotation, $r(18) = .81, p < .001$, even with trait positivity (as reported in Anderson, 1968) controlled, $pr(17) = .78, p < .001$. On moral traits relative to nonmoral ones, participants reduced just how much illusory superiority they exhibited when comparing themselves to individuals rather than to the population.

Overview of Present Studies

We conducted six additional studies that assessed our analysis of the individuation effect in illusory superiority. Study 1 aimed at replicating the finding that the largest individuation effect would be seen among morality-related traits rather than unrelated ones. Studies 2a and 2b explored whether this pattern arose because people were more likely to give individuals the "benefit of the doubt" along moral traits. But crucially, we expected that such a judgment strategy would be applied to comparisons against individuals, but not people in general (populations). Study 3 examined the causal question by manipulating whether the same set of traits were seen as more or less moral, to see if the individuation effect would arise more strongly when traits were seen as more morally-related. Studies 4 and 5 tested whether the individuation effect is triggered specifically for comparative judgments (reflecting an avoidance of relative aggrandizement on those traits seen as necessary to relational partners) or whether it emerges equivalently when judgments of the self and other are separated, either in form or time. Study 5

decomposed the individuation effect to determine whether it reflects a tempering of the self's own aggrandizement, a boost in one's assumptions about specific individuals (vs. populations), or both.

One reason that many psychological studies are underpowered is they essentially measure one data point per participant. We instead had participants offer judgments about 13 traits in two studies and at least three times that number in four studies. When studies were run in the lab, we had research assistants run as many participants as they could in one semester. When studies were run on MTurk, the funding lab split its monthly budget among studies being run that month. Across our studies, we leaned upon an average sample size of over 140 participants per condition. This allowed us to have an average of more than four thousand sets of judgments (defined by unique participant X trait combinations) in each study.

Study 1

In Study 1, participants were asked to compare themselves to a randomly selected stranger who was present in a group or to the average of all of those present. We had two main hypotheses. First, we predicted that an individuation effect would emerge—i.e., that participants would compare themselves less favorably to a randomly-selected stranger in the room than they would to the average of everyone present in the room.

Second, we predicted that the individuation effect would be driven by participants who saw a trait as particularly moral. As explained more fully in the methods section, we were able to provide a more conservative test than our reanalysis of Alicke et al. (1995, Study 1) permitted. Instead of testing whether the individuation effect emerged more strongly on moral versus non-moral traits (which might differ along many other crucial dimensions as well), we tested whether the individuation effect emerged more strongly among those who thought a particular trait was

high versus low in its moral connotation.

Method

      Participants and Design. One hundred thirty-eight undergraduates at a public American university participated as part of a longer session for which they received $15. Each participant was randomly assigned to an *individual* or *population* comparison condition.

      Procedure. Participants took part in groups ($M = 23$ participants; range = 10 to 36). Each sat at an individual computer station from which all others were visible. Participants learned they would assess themselves on 40 traits. Twenty were positive traits used in Alicke et al.'s (1995) Study 1; twenty traits were randomly sampled from positive traits catalogued by Anderson (1968). We did not include data from one trait, *aggressive*, in analyses because it was clear the trait would be seen as immoral (as opposed to moral or non-moral).

      Participants offered a *comparative rating* on all 39 traits. Participants in the *population* condition judged themselves compared to everyone else in the room. Ratings were made on 9-point scales anchored at 0 (*much less than the average of those here*) and 8 (*much more…*), with 4 explicitly labeled as well (*about the same as…*). Participants in the *individual* condition compared themselves against a specified other person in the room. As in Alicke et al. (1995), this random individual was always someone the participant did not already know. In this way, any social comparison asymmetry can be attributed to the target's status as an individual, not a friend.

      The ratings were made on a similar scale, ranging from 0 (*much less than the person*) to 8 (*much more than the person*), with 4 labeled too (*about the same as…*). Note participants in the population condition compared themselves to the average in the room, instead of positioning themselves in the distribution of the population, to make the two conditions' response scales

more comparable. To assess each trait's perceived *moral connotation*, participants rated all traits

on the extent to which they reflected one's moral character from 1(*not at all reflective of moral*

character) to 7 (*very reflective…*).

Results and Discussion

To prepare for a multi-level modeling analysis, we created a variable *target* to

differentiate the two conditions (+1: population, -1: individual). Target was a Level-1 variable

nested within trait in a random-slope, random-intercept model predicting participants'

comparative ratings. This permitted the influence of the target manipulation on the comparative

ratings to vary for different traits (random-slope), but also accounted for differences between

traits in participants' tendency to rate themselves more or less favorably than others (random-

intercept). We also included a random effect of participant. Replicating Alicke et al. (1995), we

again found participants compared themselves more favorably to the population of students (*M* =

6.17) than to a specified individual student (*M* = 5.25), $t(31,526.61) = 14.84$, $p < .001$. In other

words, participants showed an individuation effect: The better-than-average effect reduced once

the comparison standard became a specific present individual.

Does the perceived moral connotation of a trait predict the size of this individuation

effect? In proceeding, we analyzed the data using a conservative multi-level modeling approach

that took advantage of the fact that participants differed in the extent to which they saw any

given trait as high or low in its moral connotation. This method tests whether, for any specific

trait, the individuation effect (i.e., the extent to which participants compare themselves more

favorably against the population of other students instead of any specific other student) can be

traced to differences in how participants construe the moral connotation of *that particular trait*.

As one example, given that *appreciative* differs from *adventurous* in its moral connotation, but

also in numerous other ways, we would not want to conduct analyses that compare how people make judgments about their own relatively appreciativeness versus adventurousness. Instead, we wanted to control for such differences and instead capitalize on the fact that there is variability in participants' perception of just how morally relevant appreciative and adventurous are. Nonetheless, for ease of interpretation, readers can see the data depicted in this (admittedly flawed) between-trait form in Figure 1. The positive slope of the fit line illustrates that the individuation effect is greater for traits that are seen as more versus less moral. Our formal analyses avoid the just-reviewed concern.

We proceeded to test our hypothesis that for any given trait, the social comparison asymmetry would be greater for participants who saw that same trait as higher (vs. lower) in its moral connotation. We included two additional Level-1 variables: *moral connotation* (standardized for each trait) and the Target X Moral Connotation interaction. Consistent with our main hypothesis, the Target X Moral Connotation interaction was positive and significant, B = .09, $SE = 0.02$, $t(5402.88) = 4.55$, $p < .001$ (see Figure 2). Although people compared themselves 0.75 points more favorably to populations versus individuals when they saw a trait as relatively low (-1 SD) in its moral connotation, $t(16,363.02) = 10.02$, $p < .001$, this gap grew to 1.12 points among participants who saw that same trait as relatively high (+1SD) in its moral connotations was seen, $t(14,280.94) = 15.41$, $p < .001$. Stated differently, in moving from people who saw the same trait as relatively low vs. relatively high in its moral connotation, the individuation effect grew by nearly 50%.

Study 2a and 2b

Our first study found that people temper their illusory superiority over individuals (compared to groups) to the extent they see qualities as moral. We argue that this is because
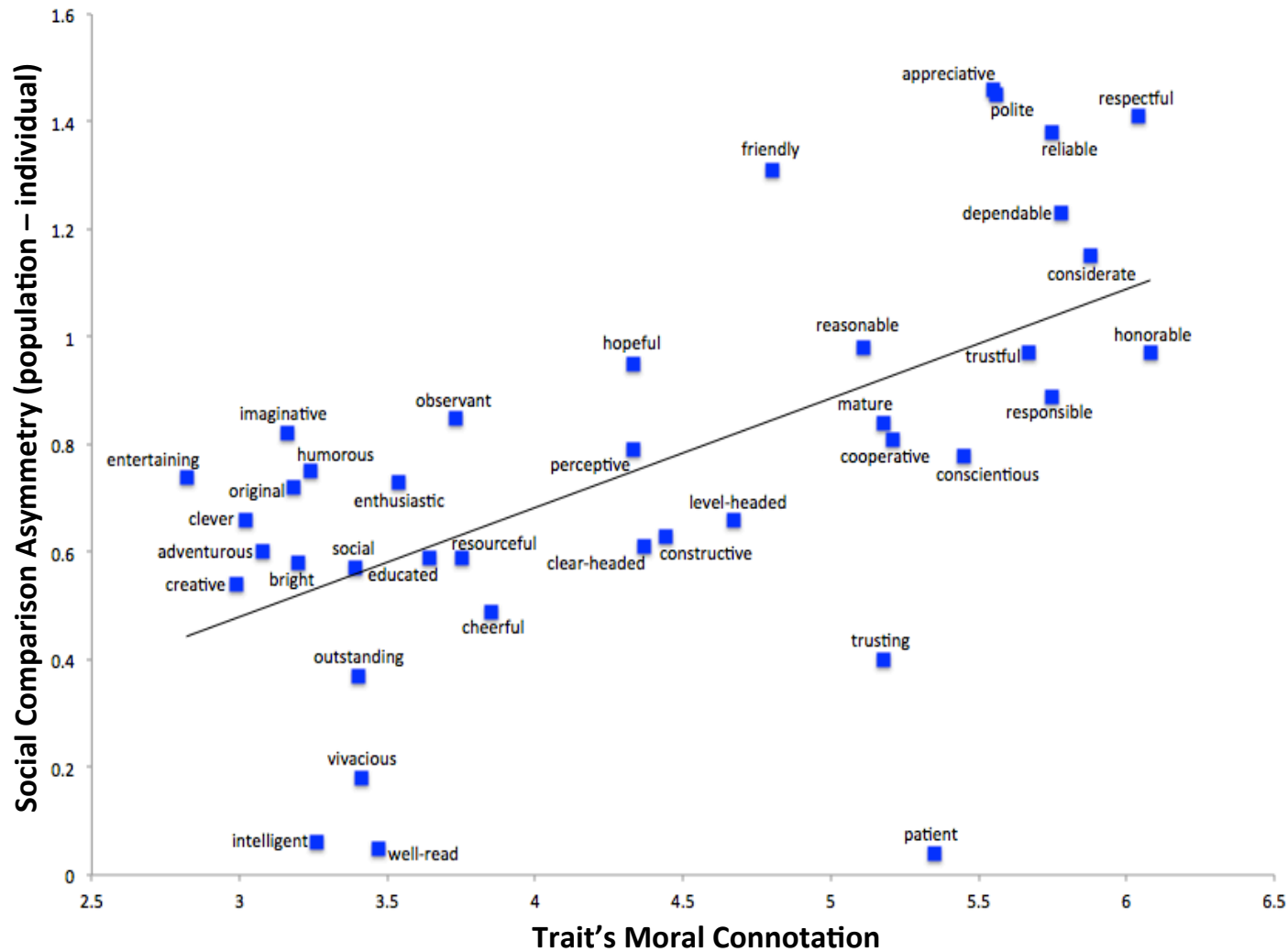
Figure 1. The average social comparison asymmetry (i.e., individuation effect) as a function of the trait's average perceived moral connotation (Study 1). The best-fit line is depicted. This visualization, although conceptually replicating our reanalysis of Alicke et al. (1995), preserves instead of (as in our main analyses) adjusts for between-trait differences in traits' moral connotation.
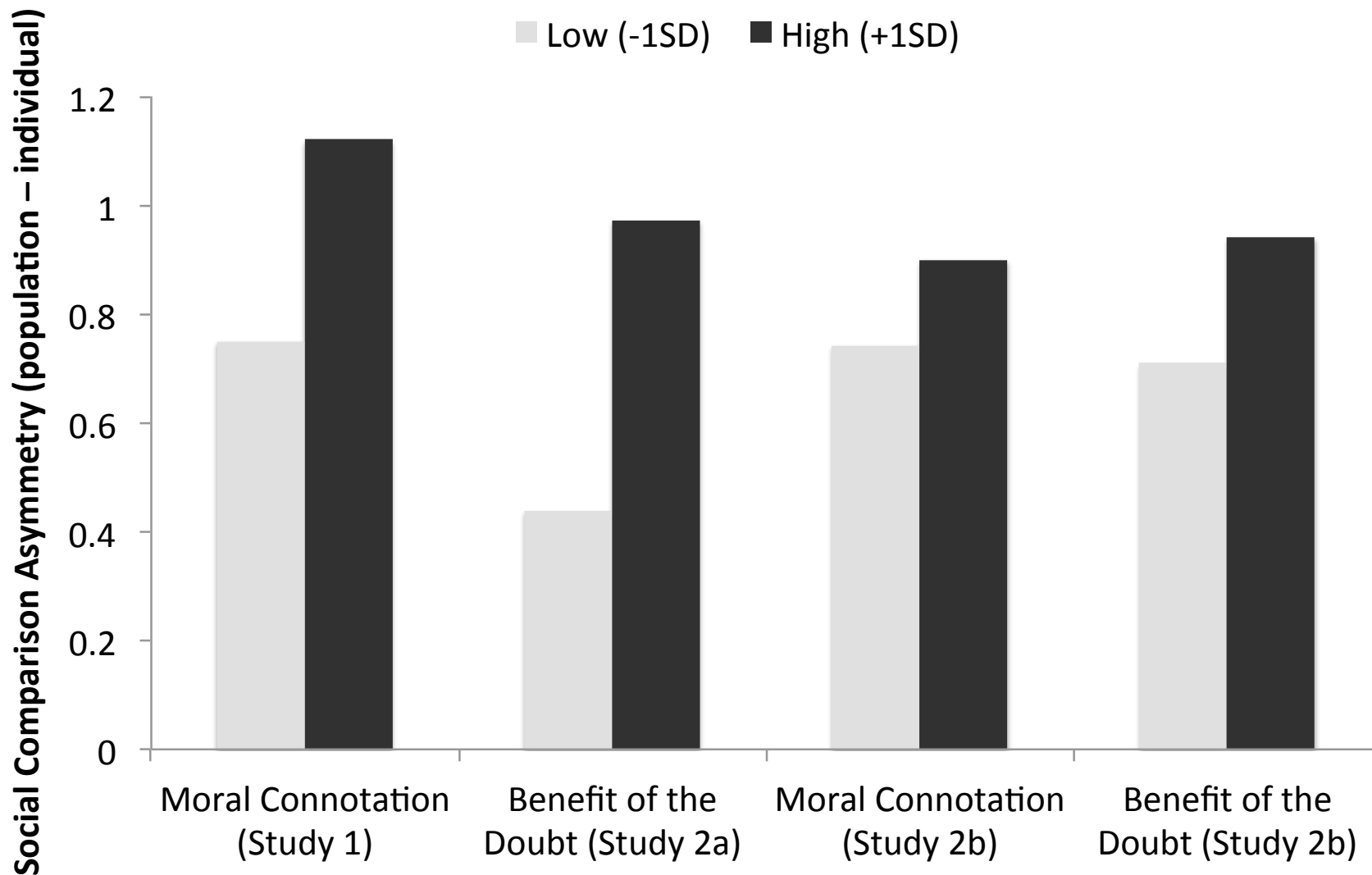
Figure 2. The model-predicted social comparison asymmetry (i.e., individuation effect) for participants who see a particular trait as relatively high (+1 SD) or relatively low (-1 SD) in its perceived moral connotation (Studies 1, 2b), or who indicate an interest in giving others a relatively high (+ 1SD) or relatively low (-1 SD) benefit of the doubt (Studies 2a-2b).

people will wish to give others the benefit of the doubt on these socially essential qualities. But crucially, this judgmental charity should be applied in contexts that parallel those in which social relationships develop. That is, the individuation effect may emerge from a desire to give specific individuals (instead of populations of people) the benefit of the doubt.

Study 2a served as an initial exploratory study of this idea, with Study 2b providing a firmer confirmatory test. Study 2a tests whether participants' indication that they would give others "the benefit of the doubt" on a trait, until proven wrong, is a strategy that could help to explain the individuation effect. More specifically, we examined whether people *selectively apply* any such benefit of the doubt strategy in comparisons against specific individuals but not people in general. Study 2b then tested whether this in part accounts for why the individuation effect is stronger when people see qualities as relatively moral (vs. non-moral).

Study 2a: Method

Participants. Seventy-nine Americans were recruited from Amazon Mechanical Turk. Each was randomly assigned to an *individual* or *population* social comparison condition.

Procedure. Participants considered themselves on 13 traits—those that were seen as most morally-average (i.e., neither clearly moral nor clearly immoral) in Study 1. Those in the *population* condition were asked to consider "all American adults" and to compare themselves on 9-point scales to the average of that population. The scale was anchored at 0 (*much less than the average of Americans*), 8 (*much more…*), with the midpoint (4) labeled as well (*about the same as…*). Those in the *individual* condition were instead shown a blurred stock image of a man or a woman. The picture had supposedly been randomly sampled from "an enormous database of pictures of randomly identified American adults." The picture was blurred to remove individuating features of the picture, so that the target's status as an individual would be salient

(see Figure 3). Participants were asked to compare themselves to this target on a similar 9-point

scale, anchored at 0 (*much less than the person*) and 8 (*much more…*). The scale midpoint (4)

was labeled as well (*about the same as…*).

Next, participants completed a new "benefit of the doubt" measure. Participants

considered the 13 traits and indicated for each whether they "tend to give people the benefit of

the doubt," assuming the best until learning otherwise, or whether they simply tend to withhold

judgment until learning more. Participants rated each trait on a 7-point scale anchored at 1 (*tend

to withhold judgment until learn more*) and 7 (*tend to just assume the best until learn otherwise*).

Note the measure did not ask about people's interest in applying the strategy to individuals or

populations in particular, so any differential application of the benefit of the doubt strategy is not

traceable to the differential applicability of the measure to the two judgment contexts.

Study 2a: Results and Discussion

We used a similar data analytic strategy to that used in Study 1. Specifically, we began by

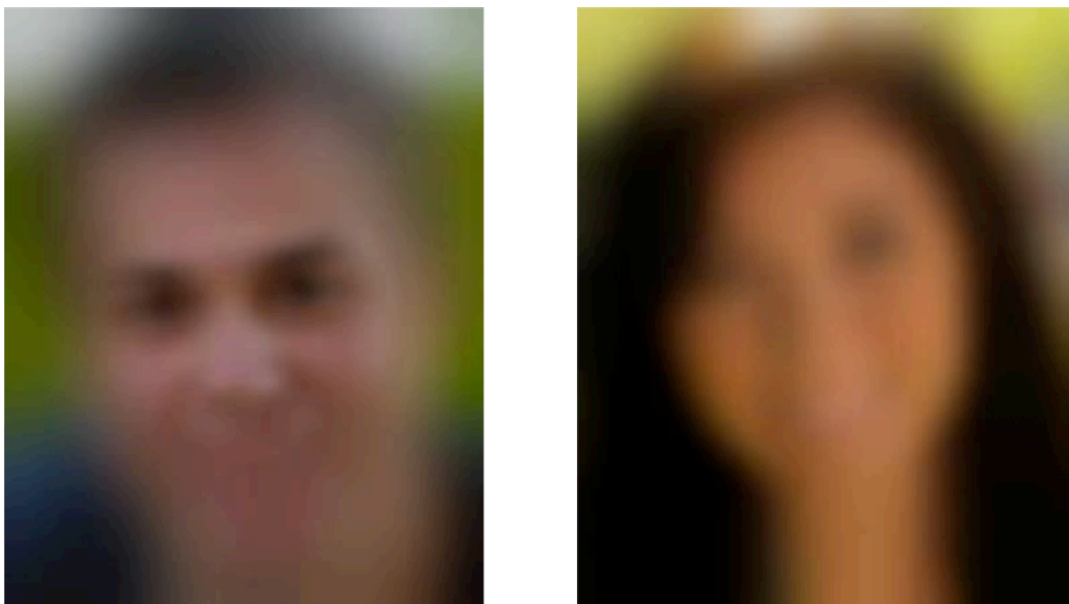defining the Level-1 variable *target*, which differentiates those participants who compared

Figure 3. Individual comparison targets.

themselves to the population of Americans (+1) or to a specific American (-1). We nested target

within trait in a random-slope, random-intercept model predicting comparative ratings. This

permitted the effect of target to vary for each of the 13 traits (random-slope), while accounting

for differences in the degree of relative self-enhancement on each trait (random-intercept).

Replicating Study 1 and Alicke et al. (1995), participants judged themselves more positively

compared to the average American ($M = 6.45$) than to an individual American ($M = 5.76$),

$t(7.42) = 5.27$, $p = .001$.

But did a differential application of participants' expressed tendency to give others the

benefit of the doubt explain the size of this individuation effect? First, we standardized the

benefit of the doubt measure within each trait pair. We included this measure as an additional

Level-1 variable nested within trait. But crucially, we added the Target X Benefit of the Doubt

interaction term, which would allow us to see whether this social strategy was applied differently

depending on the target (and thus might account for the individuation effect). Finally, we also

included a random effect of participant.

We found that the size of the individuation effect grew to the extent that participants

indicated that they gave others the benefit of the doubt on a particular trait, B $= 0.13$, $SE = 0.05$,

$t(968.72) = 2.74$, $p = .01$. As depicted in Figure 2, the social comparison asymmetry grew by

114% in moving from participants who did not express a desire to give others the benefit of the

doubt for a given trait (-1SD) to those who wished to give others the benefit of the doubt (+1SD).

Described differently, even though the benefit of the doubt measure did not ask about giving

specific individuals or people in general the benefit of the doubt, the significant interaction

suggests that the benefit of the doubt strategy was applied differently to individuals as opposed

to populations. A desire to give others the benefit of the doubt encouraged relative humility in

comparisons against specific individuals, but had not against populations. Simple-slopes analyses clarified the precise nature of this difference. When comparing oneself to all Americans, participants self-enhanced just as much regardless of their stated desired to give others the benefit of the doubt, B = -0.01, SE = 0.07, $t < 1$. In contrast, when comparing oneself to a specific (but not identifiable) American, comparative self-assessments became more humble to the extent they stated a desire to give others the benefit of the doubt on a given trait, B = -0.27, SE = 0.09, $t(44.05) = 3.17$, p = .003.

In sum, Study 2a replicated the individuation effect with a new paradigm: People compared themselves more favorably to an average American than to a blurred picture of who was said to be a randomly-selected American. But the extent to which participants said they give others the benefit of the doubt on a certain trait predicted the extent to which they showed this social comparison asymmetry. In particular, a greater willingness to give others the benefit of the doubt prompted greater humility when comparing the self against a specific individual, but had no relationship to comparative assessments against the population-at-large.

Of course, it is possible that the person—despite being represented quite minimally—still appeared to be more outstanding than the average person. Although such a tendency could distort (and even exaggerate) the size of the individuation effect, it could not account for why a self-reported desire to give others the benefit of the doubt predicted greater humility in self-assessments made in comparison to those individual targets (but not Americans in general). If anything, this concern would have predicted a different pattern—that the minimally individuating information in the blurred photograph should lead people to *depart from* (not make judgments consistent with) their decontextualized benefit of the doubt strategies.

Study 2b: Method

Participants and Design. Three hundred five Americans were recruited via Amazon's Mechanical Turk and compensated for their participation. Participants were randomly assigned to an *individual* or *population* condition.

Procedure. The procedure was identical to that used in Study 2a except for the following changes. First, participants rated themselves on the full set of 39 traits used in Study 1, instead of the reduced set of 13 traits used in Study 2a. Although we had some worries about participant fatigue on Mechanical Turk, we wanted to make sure that the findings from Study 2a generalized to a fuller range of morally and non-morally relevant traits. Second, in addition to the *benefit of the doubt* measure, participants completed the *moral connotation* measure used in Study 1. The measures were counterbalanced across participants.

Study 2b: Results and Discussion

We used a similar model to that relied upon in Study 2a. Our predictors were nested within trait in a random-slope, random-intercept model that also included a random effect of participant. Again, people compared themselves more favorably to a population ($M = 6.40$) than to a randomly-selected individual ($M = 5.55$), $t(38) = 14.52$, $p < .001$. But was this individuation effect driven by a selective application of a "benefit of the doubt" strategy when traits are seen to be more moral than non-moral?

Before proceeding to test this model within-trait (i.e., whether variation in the individuation effect for a specific trait is explained by variation in people's perceptions and strategies for that specific trait), we offer Table 1 as a trait-level (between-trait) summary of the data. Note how the social comparison asymmetry is greater for traits that are more moral in nature, and for which people indicate a desire to give others the benefit of the doubt. Like in Study 1, a significant Target X Moral Connotation interaction showed that for any given trait, the

Table 1. Indviduation Effect and Trait Characterizations by Trait (Study 2b)

Trait	Individuation Effect (Pop. – Indv.)	Moral Connotation	Benefit of the Doubt	Moral Connotation (Rank)	Benefit of the Doubt (Rank)
respectful	1.59	5.91	4.48	2	7
dependable	1.48	5.77	3.97	5	28
polite	1.47	5.47	4.92	9	2
considerate	1.35	5.73	4.49	7	6
appreciative	1.29	5.39	4.35	10	11
friendly	1.26	5.01	4.95	15	1
honorable	1.25	5.97	4.22	1	15
responsible	1.21	5.85	3.95	3	31
reliable	1.21	5.78	3.96	4	30
cooperative	1.13	5.16	4.56	13	4
reasonable	1.03	5.26	4.43	12	8
hopeful	1.02	4.51	4.41	18	9
level-headed	1.00	4.69	4.36	17	10
conscientious	0.99	5.32	4.09	11	20
trustful	0.98	5.75	3.97	6	29
bright	0.94	3.67	4.00	30	25
resourceful	0.93	3.90	3.94	22	32
perceptive	0.92	3.87	3.86	24	33
enthusiastic	0.90	3.68	4.31	29	13
humorous	0.84	3.45	4.22	33	16
observant	0.83	3.87	4.21	23	18
cheerful	0.76	4.04	4.54	21	5
educated	0.71	3.86	4.02	25	23
imaginative	0.70	3.36	3.79	36	37
constructive	0.67	4.16	4.09	20	21
well-read	0.61	3.49	3.68	32	39
trusting	0.60	5.54	3.98	8	27
intelligent	0.58	3.85	4.01	26	24
creative	0.58	3.41	3.85	35	35
mature	0.53	4.96	4.32	16	12
entertaining	0.52	3.24	4.12	37	19
clear-headed	0.49	4.36	4.25	19	14
outstanding	0.44	3.83	3.72	27	38
patient	0.44	5.15	4.22	14	17
clever	0.43	3.42	3.85	34	36
vivacious	0.39	3.17	4.08	38	22
original	0.37	3.58	4.00	31	26
social	0.23	3.70	4.61	28	3
adventurous	0.22	3.15	3.86	39	34

Note. The individuation effect reflects the average comparison of the self against the population of all Americans minus the average comparison of the self against the randomly chosen American.

individuation effect grew larger among participants who thought the trait had a higher (vs. lower) moral connotation, B = 0.04, *SE* = 0.01, *t*(11,415.97) = 2.64, *p* = .01. As depicted in Figure 2, the social comparison asymmetry grew by 21% in moving from participants who saw the same trait as relatively less moral than did others (-1 SD) to those participants who saw a trait as relatively more moral than did others (+1 SD).

In order to potentially connect this finding to Study 2a, we proceeded to test whether this individuation effect grows as people try to give others the benefit of the doubt—a strategy they may apply when comparing the self to specific individuals (but not to a population). First, we found that the more that a participant thought that a given trait was moral, the more likely the participant was to indicate a desire to give other the benefit of the doubt on the trait, B = 0.06, *SE* = 0.01, *t*(58.13) = 5.07, *p* < .001. Second, we found a significant Target X Benefit of the Doubt interaction, B = 0.06, *SE* = 0.02, *t*(11,499.17) = 3.68, *p* < .001. This replicated the key finding from Study 2a, showing that the individuation effect grew larger among participants who wished to give others, on a particular trait, the benefit of the doubt until proven wrong. The social comparison asymmetry grew by 33% in moving from participants who expressed relatively little interest (-1 SD) in giving others the benefit of the doubt to those who expressed relatively more interest (+1 SD) in giving others the benefit of the doubt (Figure 2).

Third, to test whether the social comparison asymmetry grew when traits were viewed as morally relevant *because* participants were trying to give (individual) others the benefit of the doubt, we included moral connotation, benefit of the doubt, as well as each variable's interaction with target in the same model. The Target X Benefit of the Doubt interaction continued to predict people's comparative self-assessments, B = 0.05, *SE* = 0.02 *t*(11,365.60) = 3.31, *p* = .001, but the Target X Moral Connotation interaction did as well, B = 0.04, *SE* = 0.01, *t*(11,366.36) = 2.47, *p* =

.01.

These findings are consistent with our account that seeing traits as moral prompted people to compare themselves more humbly to individuals (vs. populations) in part because of a desire to give others the benefit of the doubt when a trait was seen as more moral. Buttressing this account, simple-slopes analyses on the final model showed that although social comparisons against populations were insensitive to how much one wished to give others the benefit of the doubt on a given trait, $B = 0.02$, $SE = 0.02$, $t(192.78) = 1.02$, $p > .31$, social comparisons against specific individuals became more humble the more one wished to give others the benefit of the doubt, $B = -0.08$, $SE = 0.02$, $t(206.66) = 3.18$, $p = .002$.

Study 3

We designed Study 3 to extend on the previous studies in two ways. First, although our initial studies randomly assigned participants to make judgments comparing themselves to individuals or to populations, we leaned on natural variation between participants in terms of moral connotations they assigned to traits. In Study 3, we moved beyond this correlational approach by try to manipulate whether participants construed certain traits as more or less morally-relevant.

Second, although Studies 2a and 2b showed that for moral traits people wish to give specific individuals the benefit of the doubt, something they do in particular when they see traits as having a moral connotation, it remained vague why this is. Study 3 was designed to distinguish between two potential reasons why people might give specific others the benefit of the doubt on moral traits. Both hypotheses are rooted in functional arguments that recognize that social relationships arise with specific others instead of with people in general.

Both accounts anticipate why it may be unwise for people *not* to give specific individuals

the benefit of the doubt along moral dimensions. By the *necessity hypothesis,* people vary in the extent to which they see certain traits as absolutely necessary (vs. merely preferable) for their social relationship partners to possess. By this account, it might be socially counterproductive if people aggrandize their own relative standing on these essential qualities—at least when they are comparing themselves to specific individuals. After all, if every individual who people encountered were seen as much beneath them on qualities deemed indispensable, they would never have much motivation to engage with that individual, learn otherwise, and benefit from the relationship (Denrell, 2005; Fetchenhauer & Dunning, 2010). In contrast, people need not hold back on their desire to see the self as a superior being when considering their standing in the population at large. If traits are deemed necessary to the extent that they are seen as moral, then the necessity hypothesis might provide a nuanced answer why people give individuals the benefit of the doubt on moral traits.

In contrast, according to the *slow-to-learn hypothesis*, people appreciate that it takes more or less time to learn about different qualities of any specific person. By this alternative account, these slow-to-learn qualities may be precisely those for which people grant specific individuals the benefit of the doubt. In other words, to the extent people know they cannot appreciate the presence (or absence) of a quality in others for some time, they may be reluctant to jump to comparing themselves favorably until they have adequate information. Perhaps when an incorrect first impression of another can be corrected more quickly, people do not feel the need to temper their initial views of how much better they are on that attribute.

Pretest: Validating the Moral Connotation Manipulation

Recall that Study 2a used the 13 traits (from the set of 39 we used in our other studies) that were ambiguously moral. We exploited the middling morality of these traits by designing a

manipulation that could frame those 13 traits as relatively moral or non-moral. We explained to all participants that personality psychologists have identified 13 traits that speak to a person's moral character (moral traits), and 13 traits that give no information about a person's moral character (non-moral traits). Depending on whether we wanted to frame the ambiguous traits as moral or non-moral, we first exposed them to the 13 traits that participants in our earlier studies found least morally relevant or the 13 traits that participants in our earlier studies found most morally relevant, respectively.

More specifically, when we sought to frame the ambiguous traits as *moral*, participants were told they would first learn about 13 traits that give no information about a person's moral character. Participants saw each of the low moral connotation traits listed in parallel sentences worded, "A [TRAIT X] person could be either immoral or moral." But then, the participants saw the 13 ambiguous traits described in parallel sentences of the form, "A moral person is [TRAIT Y]." In this way, the traits were: 1) explicitly described as describing moral people, and 2) presented in contrast to 13 traits that participants would likely agree were less morally relevant.

When we sought to frame the ambiguous traits as *non-moral*, participants were exposed to a parallel, but reverse manipulation. In particular, participants were told they would first learn about 13 traits that strongly reflect a person's moral character. Participants saw each of the high moral connotation traits listed in parallel sentences worded, "A moral person is [TRAIT X.]" Next, participants saw the 13 ambiguous traits described in parallel sentences of the form, "A [TRAIT Y] person could be either immoral or moral." Thus, this parallel manipulation presented the ambiguous traits by: 1) explicitly describing them as not being informative as to a person's moral character, and 2) reinforcing this message by presenting them in contrast to 13 traits that participants were likely to greet as highly morally relevant.

We randomly assigned participants from an on-line subject pool hosted by an American university ($n = 188$) to complete either the *moral* or *non-moral* connotation framing manipulation. Participants then completed four measures that assessed the perceived *moral connotation*, *interpersonal nature*, *ambiguous definition*, and *controllability* of each trait. Moral connotation was measured with the item, "If all you knew about a person was that they were more moral than the average person, how likely do you think it is (from 0% to 100%) that they would be characterized as more __[TRAIT X]__ than average?" The perceived interpersonal measure began with a description of how traits vary in whether they describe how one interacts with other people. Participants then responded to the question, "To what extent does each trait describe one's interpersonal behavior?" on a 7-point scale anchored at 1(*not at all interpersonal*) and 7 (*very much interpersonal*), with the midpoint of 4 labeled "*as much interpersonal as not.*" The definitional ambiguity measure started with a description of how traits differ in whether "reasonable people could disagree about what behaviors constitute the trait." Participants were then asked, "For each trait, how much ambiguity is there in what the trait means?" Participants responded on a 7-point scale anchored at 1 (*clear, no ambiguity*) and 7 (*ambiguous, reasonable people may disagree*), with the midpoint of 4 labeled "*half-ambiguous, half-clear.*" The controllability measure began with a description of how traits vary in terms of whether they are essentially fixed or whether they are something over which a person has control. Controllability was measured on a 7-point scale anchored at 1(*completely controllable*) and 7 (*completely fixed*), with the midpoint of 4 labeled "*half-controllable, half-fixed,*" which we reverse-scored for the purpose of analysis.

The pretest showed that our manipulation was successful in changing the moral connotation of the traits without also altering their perceived ambiguity, controllability, or

interpersonal nature. More specifically, the moral framing led to more confidence that moral people would possess those traits ($M = 69.09\%$, $SD = 11.79\%$) than did the non-moral framing ($M = 62.78\%$, $SD = 13.66\%$), $t(186) = 3.39$, $p = .001$. In contrast, the moral framing did not lead to a perception of the traits as being more interpersonal in nature ($M = 4.73$, $SD = 0.98$) than did the non-moral framing ($M = 4.54$. $SD = 0.95$), $t(176) = 1.35$, $p > .17$. Also, the moral framing (vs. the non-moral framing) did not lead the traits to be seen as more ambiguous (Ms = 3.91 vs. 4.03) or more controllable ($M = 4.19$ vs. 4.12), ts < 1. With confidence that our manipulation changes the perceived moral connotation, but not other related properties of the traits, we proceeded to test whether experimentally manipulating the perceived moral connotation of traits would alter the size of the individuation effect.

Method

Participants and design. In order to achieve an especially large sample size, we recruited participants from both a public American university subject pool ($n = 235$) as well as Americans from Amazon Mechanical Turk ($n = 595$). These 830 participants were random assigned to one of four conditions in a 2(Target: individual or population) X 2 (Moral Framing: moral or non-moral) factorial design.

Procedure. The procedure was almost identical for those participants who took part in the lab as opposed to online. All participants began by completing either the moral or non-moral framing manipulation. At that point, participants compared themselves to a population or an individual. These judgments were made only on the 13 ambiguously moral traits on which the manipulation operated, not on the 13 traits that provided a contrasting context. For participants who took part in the lab, the population was everyone else in the room for that experimental session ($M = 18$ students per session; Range: 8 to 32) and the individual was a randomly selected

participant (specified by the experimenter uniquely for each participant in this condition) in the room. For participants who took part online, the population was everyone else in the study and the individual was the next (unseen) participant who would take part in the study.

Finally, participants completed in a counterbalanced order our two new measures designed to distinguish between two reasons why one might give specific individuals the benefit of the doubt — whether it is *necessary* that a social relationship partner have a trait, and whether one would be *slow to learn* whether another actually has a trait. We opened our description of necessity or essentialness by writing the following:

"Some positive qualities of other people are essential, such that *if we think they are absent*, we are **unlikely** to pursue a personal relationship with that person. For other qualities, even positive ones, it is easy to imagine pursuing a relationship with the person *regardless of whether they have the positive quality* in question. To what extent **would you be willing** to pursue a personal relationship with a person, even if they were NOT characterized by the trait?"

Participants then rated the 13 ambiguously moral traits on 7-point scales anchored at 1(*I would not because this quality is absolutely essential*) and 7(*I would still be willing to pursue a personal relationship*). We reverse-scored responses before entering them into analyses.

We tried to capture the notion that people are slower to learn certain traits than others with the following description:

"Traits vary in how long it takes to determine whether or not someone else has the trait in question. (For example, one **very quickly** determines if someone else is attractive). Other traits can only be detected *after a larger number of encounters*. How quickly can you detect whether someone is…?"

Participants saw the 13 ambiguously-moral traits and rated each on a 7-point scale anchored at 1(*it takes time to observe*) and 7(*very quickly on first meeting*). We reverse-scored responses prior to analyses.

Results

First, we tested whether we replicated the Alicke et al. (1995) individuation effect and whether the moral connotation manipulation moderated the size of the effect. Toward this end, we again constructed a random-slope, random-intercept model predicting participants' comparative self-ratings. We defined two Level-1 variables—target (+1 = population, -1 = individual) and moral connotation (-1 = non-moral, +1 = moral)—that were nested within trait. This permitted the effects of our two manipulations to vary for each trait (random-slope), even as we accounted for differences between traits in how much better than others participants rated themselves (random-intercept). Consistent with our hypotheses, we observed a significant Target X Moral Connotation interaction, $B = 0.11$, $SE = 0.04$, $t(740.16) = 2.58$, $p = .01$. We replicate the individuation effect when these traits are framed as moral, $B = 0.23$, $SE = 0.06$, $t(398.52) = 3.57$, $p < .001$. But when we frame the traits as non-moral, the individuation effect disappeared, $B = 0.00$, $SE = .06$, $t < 1$ (Figure 4).

Next, we proceeded to test whether the moral framing manipulation may have moderated the individuation effect by making traits seem more essential to possess or by convincing people that it would take more time to learn whether others have the trait. We used another random-slope, random-intercept model to determine whether the moral connotation manipulation changed the extent to which people saw the traits as socially necessary and slowly learnable. Both results emerged—though in the latter case, marginally. More specifically, those led to construe the ambiguously moral traits as moral (vs. non-moral) construed the traits as more
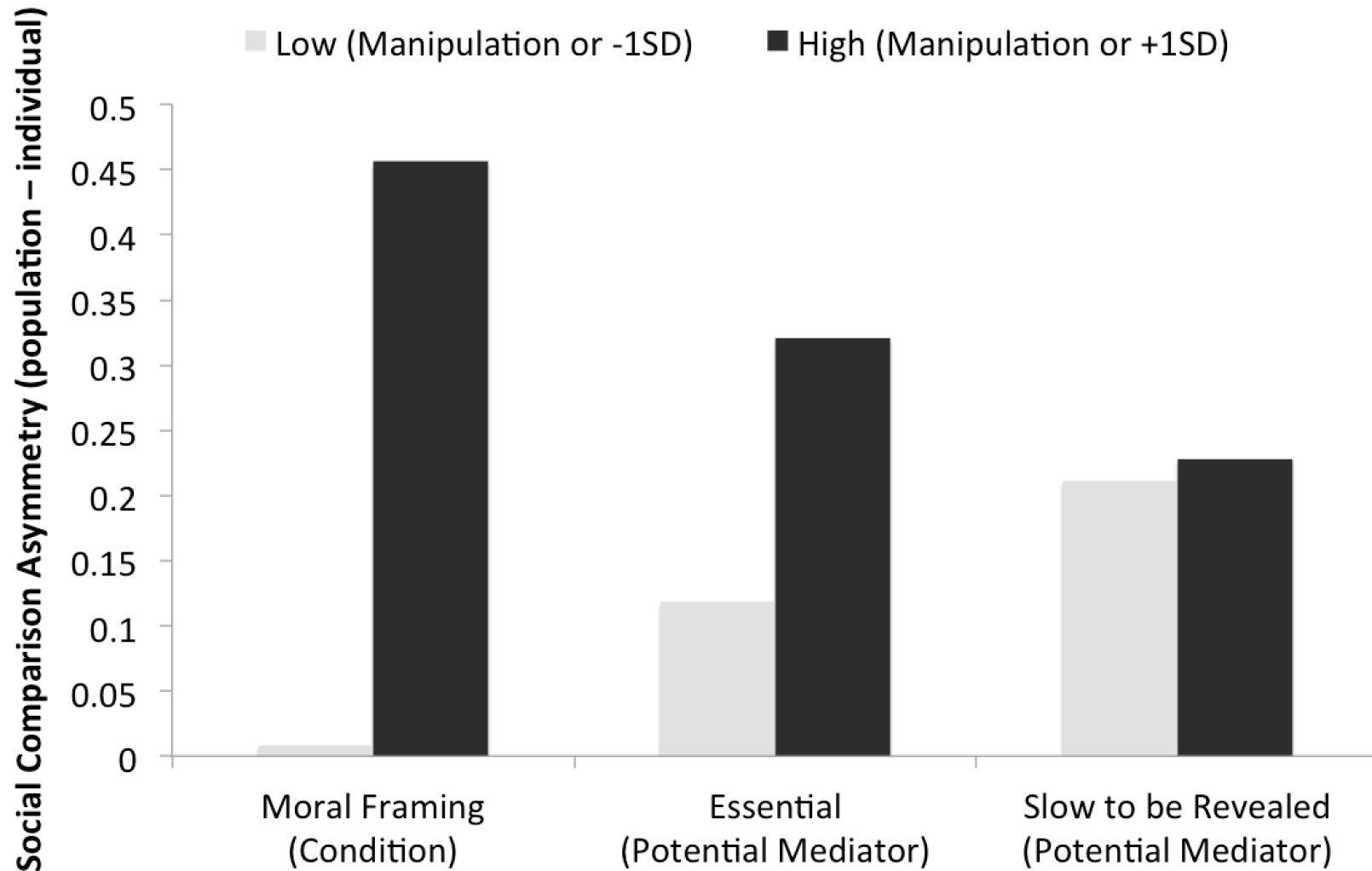
Figure 4. The model-predicted social comparison asymmetry (i.e., individuation effect) for participants who were led to perceive the traits as relatively high or low in their moral connotation, and perceived a trait to be relatively high (+1) or low (-

socially essential (Ms = 3.94 vs. 3.74), t(368.77) = 2.10, p = .04. Furthermore, framing the

ambiguous traits as moral (vs. non-moral) prompted a marginally stronger belief that they would

be slower to be revealed (Ms = 3.66 vs. 3.53), t(773.89) = 1.73, p = .08.

Might one (or both) of these consequences of the moral framing manipulation help to

clarify why seeing a trait as moral causes more of a social comparison asymmetry on the trait?

We extended our original random-slope, random-intercept model by including Level-1 main

effects of necessity and slow-to-learn (both standardized within each trait), as well as the

interaction of each with target. This model showed that the extent to which a trait was viewed as

socially necessary moderated the individuation effect, B = 0.05, SE = 0.02, t(9075.88) = 2.88, p

= .004. In moving from those who saw a particular trait as relatively less essential (-1 SD) to

relatively more essential (+1 SD), the individuation effect grew from non-significance (B = .06,

SE = .05, t(1079.66) = 1.19, p > .23) to significance (B = .16, SE = .05, t(1088.58) = 3.22, p =

.001). However, the extent to which a trait was seen as slowly learnable did not moderate this

asymmetry, B = 0.00, SE = 0.02, t < 1. In this case, moving from those who thought a particular

trait would be slow to be revealed (-1 SD) to relatively fast to be revealed (+1 SD) decreased the

individuation effect by a non-significant 8% (Figure 4). With these new terms added, the Target

X Moral Framing interaction remained significant, B = 0.11, SE = 0.04, t(714.22) = 2.51, p = .01.

In other words, that moral traits are seen to be socially necessary *partially* explained why the

moral framing manipulation caused the individuation effect to grow.

Study 4

To this point, we have had participants make single judgments that explicitly compare

themselves to a specific individual or to a population of individuals. Our reasoning has appealed

to the comparative nature of these judgments—that people are more willing to place the self

above people in general but not that far above a specific individual for traits that are seen to be

moral and thus socially necessary. But empirically, the importance of the actual social

comparison has yet to be established. That is, it remains an open question whether the

individuation effect actually requires that people make these judgments in comparative form.

By one account, what triggers the individuation effect—especially for moral traits—*is* the

comparative nature of the judgment. On the one hand, the self would like to fulfill its own desire

to see itself positively, and thus of high standing compared to others in general. On the other

hand, it may be counterproductive for the self to display the same self-aggrandizement in its

approach to specific others (Rentzsch & Schröder-Abé, 2015). After all, assuming too much of a

self-other imbalance on qualities deemed socially necessary might demotivate the self from

pursuing new social relationships. Put differently, if the self assumes it is much more reliable

than every other individual it encounters, the self may never give people a chance to prove that

they are reliable and instead counterproductively trek it alone. By this *comparative humility*

hypothesis, the individuation effect should shrink (and perhaps be driven by different

mechanisms) when the self and social judgments are no longer made in comparative form.

But by another account, the desire to give specific others the benefit of the doubt is

triggered not by the self-other comparison ("I'm way more reliable than people in general, but I

shouldn't assume I am that much more reliable than him"), but instead merely characterizes

people's patterns of self and social judgments ("I'm very reliable; he's probably somewhat

reliable; people in general aren't that reliable.") If the latter possibility is true, it suggests we

should be able to measure social judgments and self judgments at different points in time, but

then still see evidence of the same social comparison asymmetry on the comparative judgments

that these sequential judgments merely imply. That is, self-judgments minus the (individual or

population) target judgments, made minutes apart, should show the same individuation effects

for the same reasons. By this *incidental humility* hypothesis, the comparative nature of the

judgments are incidental to the effect, a mere historical artifact that the individuation effect was

first identified in the context of studying the better-than-average effect (Alicke et al., 1995).

Instead, the social comparison asymmetry may emerge equally strongly (and for the same

mechanistic reasons) regardless of whether made in comparative or sequential (separated in time)

form. Study 4 distinguishes these two hypotheses by varying whether or not participants were

prompted to judge the self and either a specific individual or the population in comparative form

(as in the previous studies) or in sequential form (as two separate judgments separated in time).

Method

      Participants and design. Four hundred fourteen Americans were recruited from

Amazon.com's Mechanical Turk and paid a nominal amount for their participation. Participants

were randomly assigned to one of four conditions in a 2(Target: population or individual) X

2(Judgment: comparative or sequential).

      Procedure. The *comparative judgment* conditions replicated the two conditions from

Study 2b, but with one minor change. We added a fortieth trait: "positive (in disposition)."

These comparative judgment conditions permitted us to test the robustness of all effects reported

in Study 2b.

      In contrast, the *sequential judgment* conditions involved making judgments about the

population (Americans) or an individual (a blurred picture of a single American) before

(unbeknownst to participants when they made the social judgments) providing assessments of

their own personality. Thus, whereas those in the comparative condition made a comparative

judgment on a scale form 0 to 8, those in the sequential judgment condition made assessments on

two 0 to 8 scales. Participants responded to the prompt "I see the average American as…"
(population) or "I see the person in the photo as…" (individual) on scales anchored at 0 (*no, not at all*) and 8 (*yes, very much so*). The midpoint of 4 was labeled "*somewhat yes, somewhat no.*"
The 40 traits appeared in a random order.

Finally, just like in Study 2b, participants completed the moral connotation and benefit of the doubt measures in a counterbalanced order.

Results and Discussion

We conducted a parallel set of tests for the comparative and sequential judgment conditions. For the sequential judgment condition, we constructed an *implied comparative judgment* by taking a difference score between self-ratings and target ratings. For each condition, we tested for an individuation effect (whether the self is judged more positively than populations as opposed to individuals), whether this social comparison asymmetry is stronger among those who see traits as more moral, and whether this exaggerated asymmetry is explained by a selective application of a "benefit of the doubt" strategy that is applied in comparisons with specific individuals as opposed to the general population. Given the ratings in the two judgment conditions are made on different scales, we proceed to analyze the conditions separately:

Comparative judgment condition. We began by conducting the same set of analyses used in Study 2b. First, we relied on a random-slope, random-intercept model (with Level-1 predictors nested within trait and with random effects of participant) to determine whether people judge themselves more positively when compared against the average of the population of Americans instead of against a specific American, as well as whether this difference was moderated by the perceived moral connotation of the trait. And indeed, participants compared themselves more favorably to the average American ($M = 6.41$) than to a specific individual

American ($M = 5.74$), B $= .34$, $SE = .07$, $t(213.71) = 4.63, p < .001$. Furthermore, participants

who saw any specific trait as higher in moral connotation showed more of this individuation

effect than did those who saw the same trait as lower in moral connotation. That is, the Target X

Moral Connotation interaction was positive and significant, B $= 0.05$, $SE = 0.02$, $t(8129.25) =$

$2.51, p = .012$. As depicted in Figure 5, the individuation effect grew by 32% in moving from

participants who saw a trait to be relatively non-moral (-1 SD) to those who saw a trait to be

relatively moral (+1 SD).

As before, the more that participants saw a trait as moral, the more they indicated it was

important to give others the benefit of the doubt on that trait until proven wrong (instead of

merely withholding judgment), B $= 0.08$, $SE = 0.01$, $t(43.43) = 5.94, p < .001$. Did an increased

interest in giving *specific* others (instead of a generalized population of others) the benefit of the

doubt explain why there was more of an individual-population social comparison asymmetry

when a particular trait was seen to be relatively high (vs. low) in its moral connotation? We

included the main effect of the benefit of the doubt variable as well as its interaction with target

to test this possibility. Replicating Study 2b, an increased interest in giving others the benefit of

the doubt on a trait enhanced the social comparison asymmetry, B $= 0.07$, $SE = 0.02$, $t(8085.08)$

$= 3.54, p < .001$, though the Target X Moral Connotation interaction remained significant as

well, B $= 0.04$, $SE = 0.02$, $t(8128.17) = 2.18, p = .030$.

We unpacked the Target X Benefit of the Doubt interaction in two ways. As depicted in

Figure 5, participants who thought that others deserved the benefit of the doubt on a particular

trait (+1 SD) showed 65% more of an individuation effect than did those who did not think

others deserved the benefit of the doubt on that same trait (-1 SD). We also tested whether an

interest in giving others the benefit of the doubt affected comparisons against populations and
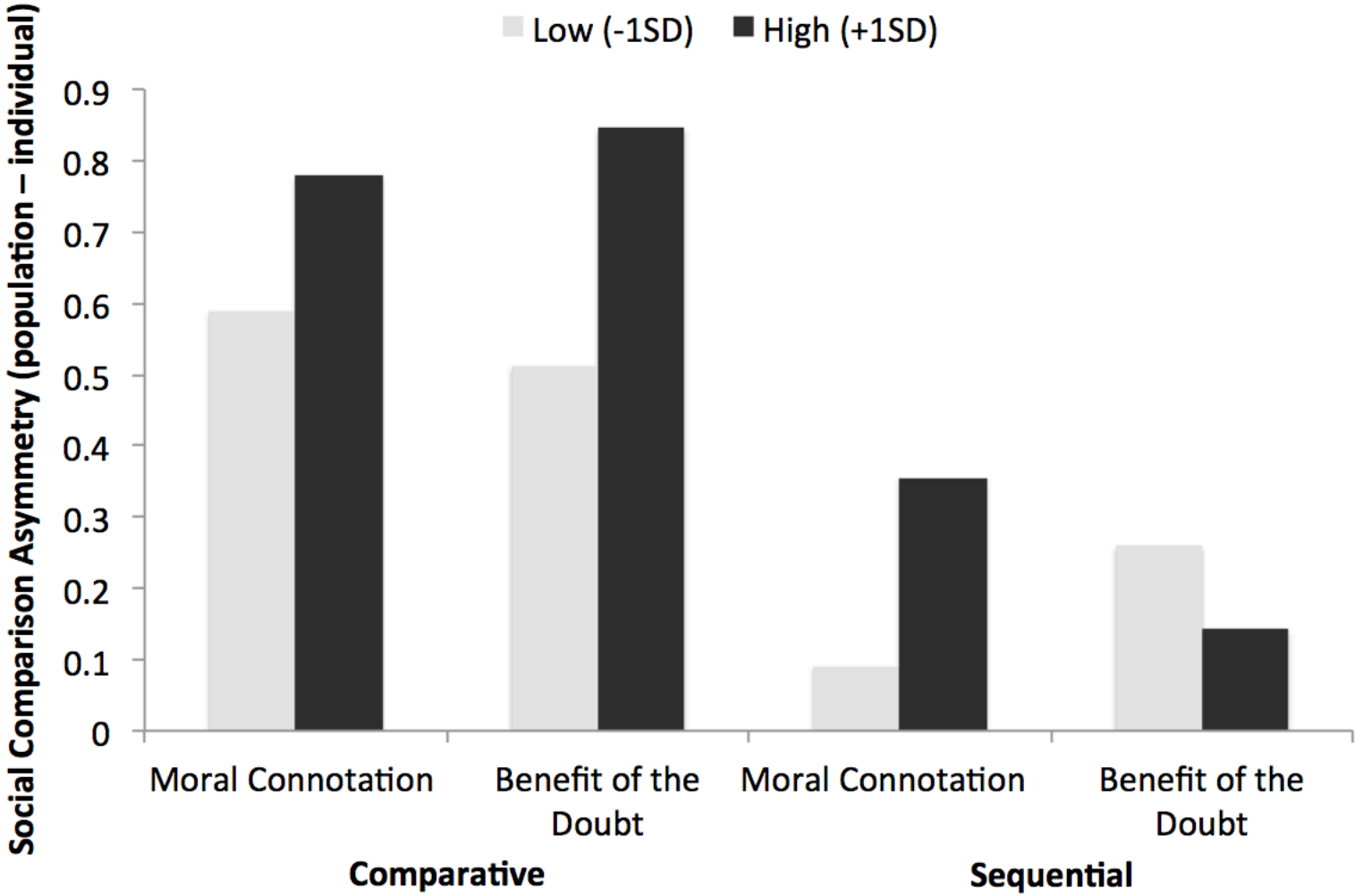
Figure 5. The model-predicted social comparison asymmetry (i.e., individuation effect) by experimental condition for

individuals differently. Those who indicated a greater interest in giving others the benefit of the doubt expressed greater comparative humility when comparing the self against a specific individual, B = -0.08, *SE* = 0.03, *t*(276.12) = -2.62, *p* = .009, but (although not predicted *a priori*) actually self-aggrandized more when comparing the self against the population, B = 0.09, *SE* = 0.03, *t*(363.55) = 3.01, *p* = .003. In summary, seeing a trait as higher in a moral connotation encouraged more humility in comparisons against a specific person (vs. people in general) in part because of an interest in giving individuals (but not populations) the benefit of the doubt on those dimensions. A significant Sobel test supported the significance of this partial mediation model, *Z* = 2.66, *p* = .01.

Sequential judgment condition. We submitted the implied comparative judgments to a similar set of analyses. But once we decoupled the self from the social judgments, we no longer observed an individuation effect, B = .11, *SE* = 0.09, *t*(212.28) = 1.16, *p* = .246. We did uncover a significant Target X Moral Connotation interaction, B = 0.06, *SE* = 0.02, *t*(7748.93) = 2.78, *p* = .005. But suggesting this effect did not emerge for the same reason it did in the comparative judgment condition, we did not uncover a significant Target X Benefit of the Doubt interaction, B = -.03, *SE* = .03, *t*(7595.32) = 1.10, *p* =.271. If anything, this interaction trended in the opposite direction. That is, in (implied) comparisons against individuals, those who expressed an interest in giving others the benefit of the doubt showed no greater (implied) humility, B = -.03, *SE* = .05, *t* < 1. It was actually (implied) comparisons against people in general in which we saw a (marginal) connection between assuming the best about others and greater (implied) humility, B = -.09, *SE* = .05, *t*(24044.63) = 1.86, *p* = .062.

Study 5

Although Study 4 showed that the individuation effect does not characterize people's self

and social judgments when made at different points in time, these results do not provide conclusive support for the *comparative humility* hypothesis. That is, the comparative and sequential judgments differed in two ways. First, and relevant to the comparative humility hypothesis, participants knowingly expressed their relative standing in the comparative judgment condition, but not in the sequential judgment condition. But second, those in the comparative judgment condition expressed their relative standing by responding to a single measure, whereas those in the sequential judgment condition responded to two measures. Study 5 provides a more conservative and thus stringent test of the comparative humility hypothesis by preserving the first but removing the second difference.

Study 5 builds on the previous studies in three main ways. First, in order to provide a stricter test of the comparative humility hypothesis, all participants expressed their self and social judgments on separate scales. We varied whether participants made their self and social judgments one after the other for each trait (concurrent condition), or made all self judgments just before or just after making all social judgments (sequential condition). Although those in the concurrent condition expressed their relative judgment on separate scales, the question prompts made clear that participants should express their own perceived relative standing on each trait through the pair of judgments. In this way, we could test whether the individuation effect emerges only when people are trying to express their relative standing on a dimension (as the comparative humility hypothesis suggests), or only when people express their relative standing through a single judgment (meaning no individuation effect should emerge in this paradigm).

Second, if the individuation effect does indeed emerge in the concurrent judgment condition (consistent with the comparative humility hypothesis), this would be informative because we would for the first time be able to see whether the individuation effect is reduced due

to a depression of self-judgments, an elevation of other-judgments, or both. In so doing, this arguably is informative about whether it is self-aggrandizement or other-derogation that prompts the individuation effect in the first place. Third, we wanted to more precisely localize the influence of moral connotation in producing the individuation effect. To isolate the influence of morality, we also measured the perceived subjectivity and controllability of each trait. Although the pretest to Study 3 showed that our experimental manipulation of the perceived morality of each trait did not influence the perceived subjectivity or controllability of those traits, we did not measure those perceived qualities in the main study. We expected our effects of interest to emerge in particular when the perceived moral connotation of traits was high, because this would encourage a selective application of the benefit of the doubt strategy when making comparisons against individuals (vs. populations).

Method

Participants and design. Nine hundred ninety-one participants—recruited from an undergraduate subject pool (n = 520) and Amazon's Mechanical Turk (n = 471)—were recruited to take part in the study. The university students participated in this and other studies in an hour-long session in exchange for course credit; AMT participants received small, monetary compensation. Participants were randomly assigned to one of four conditions in a 2(Target: population or individual) X 2(Judgment: concurrent or sequential). In an effort to identify and screen our participants who did not pay attention, we included a single, multiple-choice attention check. Seven hundred sixty-nine participants passed this check correctly: 82.5% of the undergraduate sample and 72.2% of the AMT sample.

Procedure. Participants learned that they would be rating their own personality as well as that of others on 40 different personality traits. For those in the *population* target condition, that

other was (the average) of all of the other participants who were participating in the study—

either fellow undergraduates or other AMT workers. For those in the *individual* target condition,

that other was "the last person who participated in this study." Regardless of whether participants

completed the study in the lab or at home, participants did not actually see who this person was,

meaning they had even less (than the already small amount of) individuating information than

participants in the previous studies (who saw a stranger sitting across the room or a blurred

photograph).

Those in the *sequential* condition made judgments much as did those in that condition in

Study 4. Whereas participants in Study 4 always judged the target before judging the self, for

participants in the present study we counterbalanced whether participants judged the self or the

social target first. In contrast, those in the *concurrent* judgment condition were told that they

would be asked to express a comparative judgment, but by providing two separate back-to-back

judgments. For each trial, the instructions read:

"Indicate whether you see yourself as much more TRAIT X than the previous [average]

participant, much less TRAIT X than the previous [average] participant, or about the same as

the previous [average] participant by rating both below:"

They then rated themselves and the last participant (or the average participant) in a

counterbalanced order. In all conditions, judgments were made on 0 (*no, not at all*) to 8 (*yes,*

very much so) scales, with the midpoint of 4 labeled "*somewhat yes, somewhat no*."

Next, participants completed 4 blocks of measures in a randomized order. Two of these

measures have been used in previous studies: the perceived *moral connotation* of each trait and

the endorsement of the *benefit of the doubt* strategy. Two measures were new. To assess the

perceived *subjectivity* of each trait, we explained that, "Some qualities are highly subjective, in

that people could disagree on how to define the trait. Other traits can only mean one thing." We went on to differentiate that highly subjective traits are "ambiguous, and could refer to many different behaviors," whereas those traits that are not at all subjective are "unambiguous, referring clearly to one and only one type of behavior." Participants indicated whether each of the 40 traits, presented in a random order, was subjective on a scale from 1(*not at all subjective*) to 7(*highly subjective*).

In order to capture the perceived *controllability* of each trait, we first noted that "aspects of personality differ in how *controllable* they are." We explained that some traits are highly controllable, such that "if people try or put in a little effort, they can act differently." In contrast, some traits are not at all controllable, in that "they reflect a stable aspect of a person." Participants indicated to what extent they saw each personality trait as controllable or not on a 7-point scale form 1(*not at all controllable*) to 7(*highly controllable*). As with all measures, the traits appeared in a random order.

Results and Discussion

The individuation effect on concurrent vs. sequential judgments. We began by testing for the individuation effect—i.e., whether or not people's implied comparative judgments are more favorable when made against a population as opposed to a specific (unknown) individual— and whether any such difference depends on whether such judgments are made concurrently or sequentially. Thus, for each trait for each participant, we calculated an *implied comparative judgment* by taking the self-judgment and subtracting the social target judgment. We then constructed a random-slope, random-intercept model predicting the implied comparative judgment. We defined two Level-1 variables that were nested within trait: Target (+1 = population, -1 = individual) and Judgment (+1 = concurrent, -1 = sequential). This permitted the

effects of the manipulation to vary within trait pair (random-slope) and the implied comparative judgment to vary by trait (random-intercept). We also included the Target X Judgment interaction term.

Consistent with our central hypothesis, there was a significant Target X Judgment interaction, $B = 0.10$, $SE = 0.03$, $t(765.00) = 2.84$, $p = .005$, suggesting the size of the individuation effect depended on whether judgments of the self and other were made concurrently versus sequentially. To understand whether this interaction supported the comparative humility hypothesis, we examined the effect of target separately within each judgment condition. When the two judgments were separated in time (sequential condition), participants' implied comparative judgments were actually (non-significantly) less self-aggrandizing when comparing the self against a population ($M = 0.60$) as against an unspecified individual ($M = 0.73$), $t(745.43) = -1.47$, $p = .14$. But when the two judgments were made concurrently (despite being on separate measures), participants' implied comparative judgment was more self-aggrandizing when comparing the self against a population ($M = 0.41$) than against an unspecified individual ($M = 0.16$), $t(749.30) = 2.52$, $p = .01$.

Decomposing the individuation effect into self and other judgments. By our reasoning, people have internalized a (potentially) functional bias that leads them to avoid exaggerating differences—i.e., it promotes comparative humility—between themselves and specific individuals on precisely those qualities that are crucial to interpersonal relationships. This is why the asymmetry emerges only when making direct comparative judgments, expressed either in one comparative or two distinct but concurrently offered judgments. But what we have not been able to test until now is whether the individuation effect is driven by a reduction in positivity about the self (when compared to a specific individual), an elevation in positivity about

the moral aspects of a specific individual other (when compared against the self), or both. In the context of the present study, we can ask whether the emergence of the individuation effect (when self and other judgments are made concurrently) are driven by self-humility or other-aggrandizement.

We reconducted the just-reported model twice. Instead of predicting the implied comparative judgment directly, one model predicted self-judgments and the other predicted other-judgments. Self-judgments showed a strong Target X Judgment interaction, $B = .13$, $SE = .04$, $t(765.00) = 3.67$, $p < .001$. Other-judgments did not show this same effect, $B = .04$, $SE = .04$, $t < 1$. To partially illustrate what these interactions capture, we examined how self and other judgments vary under those conditions when the individuation effect emerges—i.e., when self and other judgments are expressed concurrently. In this judgment format, self-judgments became more humble when comparing itself against an individual ($M = 6.05$) vs. a population ($M = 6.38$), $t(762.31) = 3.16$, $p = .002$. In contrast, other-judgments were equivalent regardless of whether that other was an individual ($M = 5.89$) or the population average ($M = 5.97$), $t < 1$. Thus, when the self gives specific individuals, but not populations, the benefit of the doubt, it does so by reducing its own self-aggrandizement, not by elevating its hopes for specific others. That is the self shows comparative humility.

Connecting the present findings with our previously supported mechanism. Finally, we returned to our implied comparative judgments and conducted several supplemental analyses that would allow us to connect our present findings with our past studies. For example, we found that our key Target X Judgment interaction was further moderated by the perceived moral connotation of the traits, $B = 0.02$, $SE = 0.01$, $t(30380.48) = 2.10$, $p = .04$. As would be expected based on our previous results, the hypothesized Target X Judgment interaction was especially

strong when trait moral connotation was perceived to be high (+1 SD), B = .12, *SE* = .04,

t(931.65) = 3.33, p = .001, but weakened when the moral trait connotation shrank (-1 SD), B =

.07, *SE* = .04, t(933.83) = 2.00, p = .05. In other words, the individuation effect grew stronger on

concurrently expressed judgments (compared to sequentially expressed judgments) to the extent

participants thought a particular trait was higher in its moral connotation.

Extending on our previous studies, and to more precisely identify the unique role of the

perceived moral connotation of the traits, and not their perceived subjectivity or controllability,

we tested a modified model. We added both perceived subjectivity and controllability as Level-1

variables. But also, we included new parallel 2-way and 3-way interaction terms that substituted

in subjectivity or controllability for moral connotation. Establishing the special role of perceived

moral connotation, neither the Controllability X Target X Judgment interaction, B = -.01, *SE* =

.01, nor the Subjectivity X Target X Judgment interaction, B = .01, *SE* = .01, emerged as

significant, *t*s < 1. But the Moral Connotation X Target X Judgment interaction still predicted the

implied social comparison, B = .02, *SE* = .01, t(30365.47) = 2.10, p = .04.

Furthermore, we again found that when participants viewed a particular trait as especially

high in its moral connotation, they reported a particular interest in giving others the benefit of the

doubt on that dimension, B = 0.05, *SE* = 0.01, t(50.01) = 5.29, p < .001. Next, we returned to our

earlier model that showed a significant Moral Connotation X Target X Judgment interaction. We

added the Benefit of the Doubt measure as a Level-1 variable, as well as all two-way and three-

way interaction terms (with the two manipulations: Target and Judgment). Supporting our

reasoning, the Benefit of the Doubt X Target X Judgment interaction was positive and

significant, B = .05, *SE* = .01, t(30245.01) = 4.28, p < .001. The Moral Connotation X Target X

Judgment interaction remained significant as well, B = .02, *SE* = .01, t(30392.92) = 2.04, p = .04.

This evidence is consistent with partial mediation, and once again shows that the individuation effect grows when participants see a trait as high in its moral connotation (in part) because people desire to give others the benefit of the doubt, a strategy they apply more in comparisons against individuals than in comparisons against populations.

Summary. This study identified three main findings. Figure 6 presents them in one graphic. First, the individuation effect emerges when self and other judgments are expressed concurrently (even if not on the same response scale). That is, particularly when the comparison other is an individual, the self and other bars are closer in Panel A (concurrent judgments) than in Panel B (sequential judgments.) Second, the individuation effect emerges following conditions that encourage self-humility instead of those encouraging other-aggrandizement. That is, the gap between the self and other bars shrink due to reductions in the self judgment instead of elevations in the other judgment. Third, these effects emerged more strongly to the extent qualities were seen as moral, because this encouraged a selective application of the benefit of the doubt strategy. In Panel A, one can observe the smallest individuation effect when the comparison other is an individual *and* when participants are particularly interested in giving others the benefit of the doubt.

General Discussion

More than two decades ago, Alicke et al. (1995) found that people's tendency to see themselves as better than average was diminished when they compared themselves against specific comparison others. An explanation for this influential finding has long remained elusive. In the present paper, we claim to have solved this mystery. In so doing, we have demonstrated a connection between social representations, morality, humility, and beliefs about what qualities
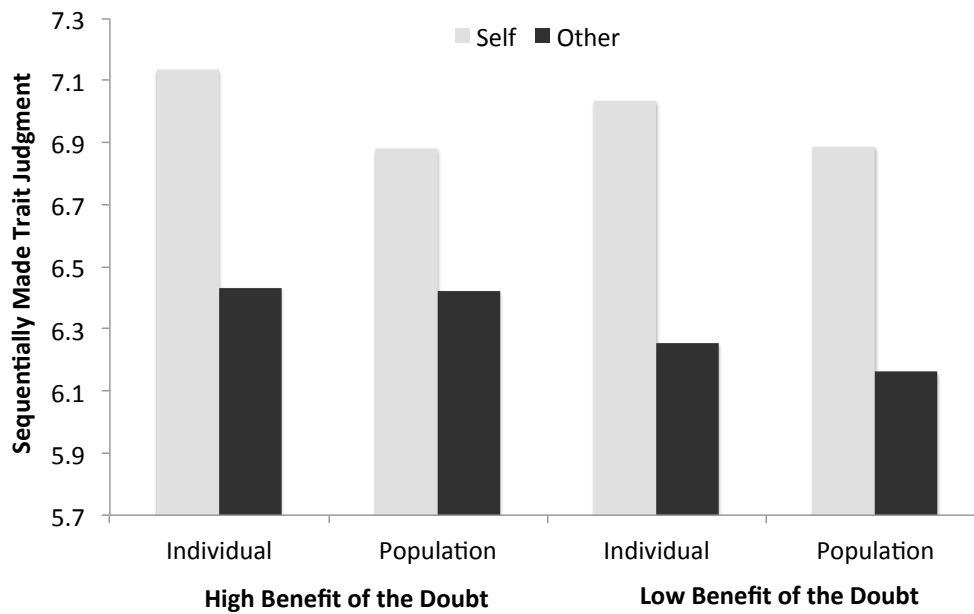
(A)
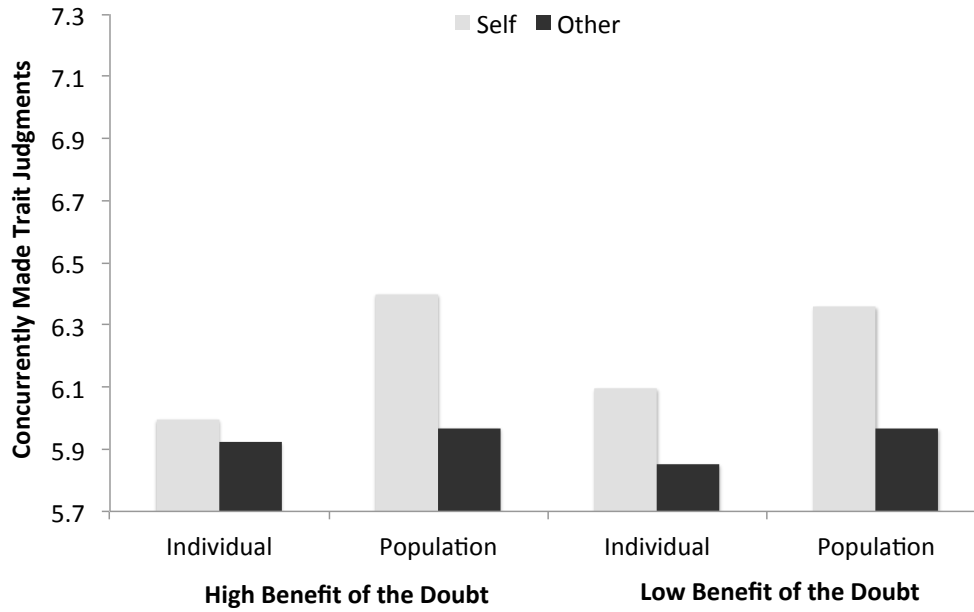
(B)

Figure 6. Self- and other-judgments when that other is the previous participant (Individual) or the average of all participants (Population) for participants who are especially likely to indicate for a given trait a willingness to give others the benefit of the doubt (+1 SD) or not (-1 SD) for participants who (Panel A) made such judgments concurrently or (Panel B) sequentially (Study 5).

are essential to social relationships. We organize the contributions of the present work into five points, which we discuss next.

First, we found robust evidence for the individuation effect across a variety of paradigms. People judged themselves more favorably when comparing themselves to a population of individuals as opposed to specific individuals from that population. Such an effect was observed when participants compared themselves to a randomly chosen stranger in a room of fellow participants and when American participants compared themselves to Americans in general or a blurred image of a randomly selected American. Furthermore, we replicated these effects when all individuating cues were hidden—i.e., when people compared themselves to the previous or next (never seen) participant.

Second, we found that the individuation effect could be tied to the perceived morality of the traits being considered. Our reanalysis of Alicke et al. (1995, Study 1) found that the size of the individuation effect (the degree to which the self became more humble when comparing itself against a person instead of a population) strongly correlated with the perceived moral connotation (above and beyond the perceived positivity) of each trait. In our main studies, we measured the individuation effect and the perceived moral connotation of traits using the same participants. This allowed us to see that even among participants considering the same trait, they showed a reduced individuation effect to the extent that they considered *that particular trait* to be higher or lower in its perceived morality (Studies 1, 2b, 4-5).

We also established the causal role that perceived moral connotations play: Prompting people to see qualities as more or less morally relevant produced relatively stronger and weaker individuation effects, respectively (Study 3). A pretest to this experimental study showed that by manipulating the perceived morality of traits, we did not also influence the ambiguity,

controllability, or interpersonal nature of the traits. In Study 5, we were able to show the unique effect of perceived moral connotation even when controlling for (the non-significant influence of) the traits' perceived subjectivity and controllability.

Third, we offered support for our argument that people become more humble in their comparative judgments against individuals because of a desire to offer others the benefit of the doubt in some such comparisons (Studies 2a-2b, 4-5). More specific tests showed that offering the benefit of the doubt was a generally endorsed strategy that encouraged humility in comparisons against individual people, but not people in general. Also central to our overall argument, the more that traits were perceived to be moral in their connotation, the more people endorsed giving others the benefit of the doubt on such traits (Studies 2b, 4). In other words, the perceived moral connotation of traits encourages an individuation effect because it prompts people to selectively apply a strategy of giving others the benefit of the doubt in comparisons against specific individuals but not people in general.

Fourth, we sought to probe and elaborate on our functional account by disentangling two reasons why the perceived moral connotation of traits encourages the individuation effect. That is, why do people wish to avoid expressing clear comparative superiority over specific others on moral dimensions? Leading people to see the same traits as moral made those traits seem: 1) more essential for social partners to possess and 2) (marginally) slower to be revealed. But only the former quality—the enhanced social necessity of moral qualities—explained the individuation effect (Study 3). That is, the perception that moral qualities are those that individual social investment partners have to possess is what explained the fact that individuals were reluctant to proclaim their superiority to specific individuals (as opposed to people in general) on these essential dimensions.

Fifth, we established that the comparative nature of the judgment is essential. That is, people are reluctant to express their own superiority to specific individuals on those qualities they deem interpersonally necessary. When self-judgments and social judgments were separated in time, the individuation effect was eliminated (Studies 4-5). That said, what encouraged this relative humility was not expressing one's relative standing in a single judgment, but instead that one aimed to communicate one's relative standing, even if such an expression relied on a pair of concurrent judgments. More specifically, participants showed an individuation effect when judging the self and another back-to-back, but not when making those two judgments at two separate points in time (Study 5). Further reinforcing that these effects reflect comparative humility instead of mere aggrandizement of others, this final study showed the individuation effect reflects a reduction in self-aggrandizement instead of a tendency to glorify an unknown other. In other words, the self's interest in giving others the benefit of the doubt is actually enacted through the self's descending from its own high horse.

By showing that the individuation effect emerges to the extent judges view traits as moral, and thus wish to give others the benefit of the doubt, have we actually explained the root cause of the individuation effect? Or have we merely shown what moderates it? We wish to emphasize that these two possibilities are not mutually exclusive. That is, by arguing that an effect emerges because of a precipitating condition, one should test this mechanistic logic by determining whether the effect emerges more strongly when the hypothesized precursor is present as opposed to absent. By tracing the pathway by which that proposed precursor (i.e., perceived moral connotation) feeds forward to ultimately produce the individuation effect (i.e., by encouraging one to assume the best about others, which is then selectively applied in comparisons against individuals as opposed to populations), we flesh out the mechanistic

pathway. That said, the individuation effect—especially given its robustness—is likely multiply determined, and we leave it for future research to uncover additional mechanisms.

Relation to Previous Work

In explaining the individuation effect—the tendency to self-enhance more in judgments against individuals than populations—how useful is it to draw parallels with what is known to produce or moderate self-enhancement as a whole? For example, Alicke (1985) showed that the controllability of dimensions moderates how much people self-enhance on them. Closer to the present work, Allison, Messick, and Goethals (1989) found that there was more self-enhancement on moral dimensions than there was on intellectual dimensions. This was because there was more interpretational ambiguity in determining whether one had behaved morally, whereas intellectual behaviors were more specific and objective. But note that qualities that influence the degree of self-enhancement—merely logically—need not explain the individuation effect. This is because the individuation effect is itself a comparison of two self-enhancing effects: how much self-enhancement is seen in comparisons against a population and in comparisons against a specific individual. That said, we did also find that neither trait ambiguity nor controllability could account for our effects. In other words, the present and previous research differ not merely conceptually, but in their empirical signatures as well.

Readers may note some irony that we opened this paper by noting that the self-enhancement literature has largely neglected the importance of the "other" in self-enhancement. Our final study demonstrated that giving others the benefit of the doubt actually manifests as a reduction in the positivity of self-judgments—i.e., comparative humility. But we wish to bring up a key distinction--whether self-enhancement *neglects* considerations of the comparison other or involves *modified* judgments of that other. Much previous work has taken the former position

by questioning whether the better-than-average effects involve comparisons at all. From this perspective, people may see themselves as better than other because they weight the focal target (the self and the generally positive views in how it sees itself) in such comparisons and essentially neglect the comparison standard (Gruenther & Alicke, 2010; Kruger, Windschitl, Burrus, Fessel, & Chambers, 2008). This explains why the self compares itself less favorably to others on particularly difficult tasks (Kruger, 1999).

Although the present research found that the individuation effect involves modifications of self instead of social perceptions, we also found that the psychology underlying this effect contradicted the idea that comparative judgments neglect social considerations. First, the individuation effect itself attests to the importance of the comparison other. Second, it was a wish to give others the benefit of the doubt on socially essential qualities that kept the self's aggrandizement in check. In other words, it was a concern with not disparaging the comparison other that the self tempered its own self-perceptions. In order to achieve the goal of comparative humility, it may be simpler to tinker with the representations of a rich target like the self than that of an informationally improverished social target like the unknown comparison other.

In explaining the individuation effect, we have made something of a functional argument—that the self is best served not seeing itself as vastly superior to specific individuals on qualities that it would be essential for individuals to possess (Dunning, 2017; Dunning et al., 2016). It is notable that our participants selectively applied an "assume the best" strategy even when they did not plan to interact with (Studies 1 and 3), actually meet (Studies 2a-2b, 4), or have the ability to even know the identity of (Study 5) the comparison other. Previous research has identified how people's attitudes and impressions shift in advance of interactions with specific others—for example, by tuning their own attitudes to match their partner's (Higgins &

Rholes, 1978; Ledgerwood, Trope, & Chaiken, 2010; McCann & Higgins, 1992), increasing

expectations of liking (Darley & Berscheid, 1967), or assuming their partner will complement

their own levels of dominance (Tiedens, Unzueta, & Young, 2007). One possibility is that the

presently documented effects would be even stronger when interaction is imminent, meaning that

the target is being considered for social investment. But another possibility is that this bias

mostly serves to encourage people to keep an open mind toward any specific individuals (like

those in the present research), not that it comes online just before any one specific encounter.

Our argument has been premised on the idea that social interactions occur with specific

individuals. But people do not interact with individuals merely in isolation, but with individuals

who are part of specific small groups. Anderson, Srivastava, Beer, Spataro, and Chatman (2006)

found that group members who have a misguidedly positive understandings of their status within

groups are disruptive, create more discord, and are ultimately less accepted by the group. One

possibility then is that we would observe a similar reduction in moral self-enhancement when

people compare themselves to small groups of indentifiable individuals—a collection of people

one may consider investing in socially. By Anderson et al.'s (2006) account, people's failure to

know their place can prove disruptive to the smooth functioning of the group. By the present

account, the negative of moral self-enhancement is it may discourage the formation of small

groups to begin with.

In considering what underlies people's desire to give others the benefit of the doubt, it is

natural to ask whether we should expect individual differences in the individuation effect. One

intriguing possibility is that individuals who have become less open to expanding their social

networks—such as older individuals, who tend to become more present-focused and less

interested in new social investments (Löckenhoff & Carstensen, 2004)—may show smaller

individuation effects. Johnston (2016) found that as seniors aged, they judged individual faces to be less trustworthy. Although this research did not examine participants' beliefs about the trustworthiness of people in general, the finding that older seniors were less inclined to give individuals' moral character the benefit of the doubt is consistent with our reasoning.

Furthermore, cultural differences in orientations toward humility may influence the individuation effect. Humility itself is prized more in collectivist cultures (Ho & Chiu, 1994; Triandis, 1989). On the one hand, this might encourage more of an individuation effect. That is, if comparing the self favorably to specific individuals feels less humble than comparing the self favorably to people in general, then the individuation effect may be a socially conditioned pattern of self and social judgment. On the other hand, these strong norms toward modesty may lead self-enhancement to be depressed in comparisons against both populations and individuals (Kurman, 2003).

Are the present findings at odds with previous research that has characterized social perception as cynical? For example, people show a tendency to become highly suspicious about the motives behind others' seemingly selfless behavior (Critcher & Dunning, 2011; Fein, 1996). Remember though that our participants reported assuming the best "until they learned otherwise." It may not take long or much for people to feel they have learned otherwise.

Recent evidence suggests that knowingly making oneself vulnerable to others who do not necessarily have a history of trustworthy behavior can bring out prosocial tendencies in those others (Bruni & Tufano, 2017). In this way, we believe that people entering new social situations may initially err toward optimism and thus reap (or even elicit) the benefits of testing out many possible social partners, but then err toward cynicism to avoid the costly mistake of continued, misdirected social investment. Though as Study 5 demonstrated, such optimism about

others manifests more as a deflation of self-perceptions instead of a direct inflation of social

perceptions. By analogy, homeowners' ability to see the shortcomings in their own houses may

make them feel it is worthwhile to search for what else is out there on the market, even as those

homebuyers may shift toward incredibly high standards before making an offer on a new home.

This shift from optimistic exploration to critical evaluation may be a general, adaptive dynamic.

Conclusion

In summary, the present results help to resolve a longstanding mystery in the social

comparison literature. More generally, these findings may help in explaining how people can

satisfy their motivations to maintain an elevated sense of self-worth (Critcher et al., 2011;

Dunning et al., 1989; Guenther & Alicke, 2010; Taylor & Brown, 1988) without undermining

the self's need to forge social bonds of trust and cooperation with specific others (e.g., Williams

& Nida, 2011). That is, people can see themselves as a member of the elite few when considering

how they stack up against a population, but then temper their sense of moral exceptionalism in

considering specific others. If finding others of sufficient moral character is what people aim for

in seeking out new ties, people should benefit from not beginning their search with an inflated

understanding of what they themselves bring to the table.

References

Abele, A. E., Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*, 751-763.

Alicke, M. D., & Govorun, O. (2005). The Better-Than-Average Effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *Studies in self and identity. The Self in Social Judgment* (pp. 85-106). New York: Psychology Press.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology, 68*, 804-825.

Alicke, M. D., & Sedikides, C. (2011). *Handbook of self-enhancement and self-protection.* New York: Guilford Press.

Allison, S T., Messick, D. M., & Goethals, G. R. (1989). On being better but not smarter than others: The Muhammad Ali effect. *Social Cognition, 7*, 275-296.

Anderson, N. H. (1968). Likeableness ratings of 555 personality-trait words. *Journal of Personality and Social Psychology*, *9*, 272-279.

Anderson, C., Srivastava, S., Beer, J. S., Spataro, S. E., & Chatman, J. A. (2006). Knowing your place: Self-perceptions of status in face-to-face groups. *Journal of Personality and Social Psychology, 91*, 1094-1110.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin, 99*, 229–246.

Brown, J. D., & Han, A. (2012). My better half: Partner enhancement as self-enhancement. *Social Psychological and Personality Science*, *3*, 479-486.

Bruni, L., & Tufano, F. (2017). The value of vulnerability: The transformative capacity of risky trust. *Judgment and Decision Making*, *12*, 408-414.

Colvin, C. R., Block, J., & Funder, D. C. (1995). Over positive self-evaluations and personality: Negative implications for mental health. *Journal of Personality and Social Psychology, 68*, 1152-1162.

Cosmides, L, Tooby, J., Fiddick, L. & Bryant, G. (2005). Detecting cheaters. *Trends in Cognitive Sciences, 9*, 505-506.

Crary, W. G. (1966). Reactions to incongruent self-experiences. *Journal of Consulting Psychology, 30*, 246 –252.

Critcher, C. R., & Dunning, D. (2011). No good deed goes unquestioned: Cynical reconstruals maintain belief in the power of self-interest. *Journal of Experimental Social Psychology, 47*, 1207-1213.

Critcher, C. R., Helzer, E. G., & Dunning, D. (2011). Self-Enhancement via redefinition: Defining social concepts to ensure positive views of self. In M. D. Alicke, & C. Sedikides (Eds.), *Handbook of self-enhancement and self-protection* (pp. 69-91). New York, NY: The Guilford Press.

Darley , J. M., & Berscheid, E. Y. (1967). Increased liking as a result of the anticipation of personal contact. *Human Relations, 20*, 29-40.

Denrell, J. (2005). Why most people disapprove of me: Experience sampling in impression formation. *Psychological Review, 112*, 951-978.

Dunning, D. (2005). Self-insight: Roadblocks and detours on the path to knowing thyself. New York and Hove: Psychology Press.

Dunning, D. (2017). Normative goals and the regulation of social behavior: The case of respect. *Motivation & Emotion, 41*, 285-293.

Dunning, D, Anderson, J. E., Schlösser, T., & Ehlebracht, D. (2014). Trust at zero acquaintance:

More a matter of respect than expectation of reward. *Journal of Personality and Social Psychology, 107*, 122-141.

Dunning, D., Fetchenhauer, D., & Schlösser, T. (2016). The psychology of respect: A case study of how behavioral norms regulate human action. In A. Elliot (Ed.), *Advances in motivation science* (vol. 3; pp. 1-34). New York: Elsevier.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*, 69-106.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology, 57*, 1082-1090.

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D. & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*, 98-121.

Fein, S. (1996). Effects of suspicion on attributional thinking and the correspondence bias. *Journal of Personality and Social Psychology, 70*, 1164–1184.

Fein, S., McCloskey, A. L., & Tomlinson, T M. (1997). Can the jury disregard that information? The use of suspicion to reduce the prejudicial effects of pretrial publicity and inadmissible testimony. *Personality and Social Psychology Bulletin, 23*, 1215-1226.

Fetchenhauer, D., & Dunning, D. (2009). Do people trust too much or too little? *Journal of Economic Psychology, 30*, 263–276.

Fetchenhauer, D., & Dunning, D. (2010). Why so cynical? Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others. *Psychological Science, 21*, 189-193.

Fetchenhauer, D., & Dunning, D. (2012). Betrayal aversion versus principled trustfulness—How

to explain risk avoidance and risky choices in trust games. *Journal of Economic Behavior

and Organization, 81*, 534-541.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person

perception and evaluation. *Journal of Personality and Social Psychology, 106*, 148-168.

Guenther, C. L., & Alicke, M. D. (2010). Deconstructing the better-than-average effect. *Journal

of Personality and Social Psychology, 99,* 755-770.

Haidt, J., & Kesebir, S. (2010). Morality. In S. Fiske, D. Gilbert, & G. Lindzey (Eds.), *Handbook

of Social Psychology, 5th Edition* (pp. 797-832). Hoboken, NJ: Wiley.

Hartley, A. G., Furr, R. M., Jayawickreme, E., Helzer, E. G., Velasquez, & Fleeson, W. (2016).

Morality's centrality to interpersonal evaluation. *Social Psychological and Personality

Science, 7*, 648-657.

Helzer, E. G., & Critcher, C. R. (in press). What do we evaluate when we evaluate moral

character? For K. Gray & J. Graham (Eds.), *Atlas of moral psychology.* Guilford Press.

Higgins, E. T., & Rholes, W. S. (1978). "Saying is believing": Effects of message modification

on memory and liking for the person described. *Journal of Experimental Social Psychology,

14*, 363-378.

Hilbert, M. (2012). Toward a synthesis of cognitive biases: How noisy information processing

can bias human decision making. *Psychological Bulletin, 138*, 211-237.

Hitchcock, C. H., Rees, C., & Dalgleish, T. (2017). The devil's in the details: Accessibility of

specific personal memories supports rose-tinted self-generalizations in mental health and

toxic self-generalizations in clinical depression. *Journal of Experimental Psychology:

General, 146*, 1286-1295.

Ho, D. Y., & Chiu, C. (1994). Component ideas of individualism, collectivism, and social

    organization: An application in the study of Chinese culture. In U. Kim, H. C. Triandis, C.

    Kagiticibasi, S. C. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method,*

    and applications (pp. 123-136). Thousand Oaks, CA: Sage.

Johnston, C. (2016). *Physiognomics: A comparative study of Chinese-American and Caucasian-*

    American senior females' use of nonverbal cues to determine trustworthiness. Unpublished

    doctoral dissertation, Grand Canyon University.

Klein, W. M. P., & Goethals, G. R. (2002). Social reality and self-construction: A case of

    "bounded irrationality?" *Basic and Applied Social Psychology, 24*, 105-114.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature

    of comparative ability judgments. *Journal of Personality and Social Psychology, 77*, 221-

    232.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing

    one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social*

    Psychology, 77, 1121-1134.

Kruger, J., & Gilovich, T. (1999). "Naïve cynicism" in everyday theories of responsibility

    assessment: On biased assumptions of bias. *Journal of Personality and Social Psychology,*

    76, 743-753.

Kruger, J., Windschitl, P., D., Burrus, J., Fessel, F., & Chambers, J. R. (2008). The rational side

    of egocentrism in social comparisons. *Journal of Experimental Social Psychology, 44*, 220-

    232.

Kurman, J. (2003). Why is self-enhancement low in certain collectivist cultures? An

investigation of two competing explanations. *Journal of Cross-Cultural Psychology, 34*, 496-

510.

Kwan, V. S., John, O. P., Robins, R. W., & Kuang, L. L. (2008). Conceptualizing and assessing

self-enhancement bias: A computational approach. *Journal of Personality and Social

Psychology*, *94*, 1062-1077.

Ledgerwood, A., Trope, Y., & Chaiken, S. (2010). Flexibility now, consistency later:

Psychological distance and construal shape evaluative responding. *Journal of Personality

and Social Psychology*, *99,* 32-51.

Löckenhoff, C. E., & Carstensen, L. L. (2004). Socioemotional selectivity theory, aging, and

health: The increasingly delicate balance between regulating emotions and making tough

choices. *Journal of Personality, 72*, 1395-1424.

McCann, C.D, & Higgins, E. T. (1992). Personal and contextual factors in communication: A

review of the "communication game". In G. R. Semin & K. Fiedler (Eds.), Language,

interaction and social cognition (pp. 144–171). London: Sage.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review,

115*, 502-517.

Newman, G. E. (2014). Tainted altruism: When doing some good is evaluated as worse than

doing no good at all. *Psychological Science, 25*, 648-655.

Paulhus, D. L. (1998). Interpersonal and intrapsychic adaptiveness of trait self-enhancement: A

mixed blessing? *Journal of Personality and Social Psychology, 74*, 1197-1208.

Ratner & Miller, (2001). The norm of self-interest and its effects on social action. *Journal of

Personality and Social Psychology*, *81*, 5-16.

Rentzsch, K., & Schröder-Abé, M. (2015). Self-enhancement 2.0: An integrated approach to measuring dyadic self-enhancement at two levels. *Social Psychological and Personality Science*, *6*, 251-258.

Sweeny, K., & Krizan, Z. (2013). Sobering up: A quantitative review of temporal declines in expectations. *Psychological Bulletin, 139*, 702-724.

Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, *103*, 193-210.

Tiedens, L. Z., Unzueta, M. M., & Young, M. J. (2007). An unconscious desire for hierarchy? The motivated perception of dominance complementarity in task partners. *Journal of Personality and Social Psychology, 93*, 402-414.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review, 96*, 506-520.

Williams, K. D., & Nida, S. A. (2011). Ostracism: Consequences and coping. *Current Directions in Psychological Science, 20*, 71-75.

Wojciszke, B. (1994). Inferring interpersonal attitudes: Hypotheses and the information-gathering process. *European Journal of Social Psychology, 24*, 383-401.

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin, 24*, 1245–1257.