9-1989

On Being Better but not Smarter than Others: The Muhammad Ali Effect

Scott T. Allison
University of Richmond, sallison@richmond.edu

George R. Goethals
University of Richmond, ggoethal@richmond.edu

David M. Messick

Recommended Citation

Allison, Scott T., David M. Messick and George R. Goethals. "On Being Better but not Smarter than Others: The Muhammad Ali Effect." *Social Cognition*, 7, no.3 (September 1989): 275 - 295.

Social Cognition, Vol. 7, No. 3, 1989, pp. 275–296

ON BEING BETTER BUT NOT SMARTER THAN OTHERS: THE MUHAMMAD ALI EFFECT

SCOTT T. ALLISON
University of Richmond

DAVID M. MESSICK
University of California, Santa Barbara

GEORGE R. GOETHALS
Williams College

Past research suggests that people believe that they perform socially desirable behaviors more frequently and socially undesirable behaviors less frequently than others (Goethals, 1986; Messick, Bloom, Boldizar, & Samuelson, 1985). The present research examined whether this perception also characterizes people's thinking about intelligent and unintelligent behaviors. In Study 1, subjects wrote lists of behaviors that they or others did. Subjects indicated that they performed more good and intelligent behaviors and fewer bad and unintelligent behaviors than others, although the magnitude of these differences was greater for good and bad acts than for intelligent and unintelligent ones. In Study 2, a different group of subjects judged the frequency with which the behaviors generated in the first study occur. While self-ascribed good behaviors were rated as occurring more frequently than the good acts of others, self-ascribed intelligent behaviors were not judged as more frequent than the intelligent acts of others. Study 3 replicated this effect using a different methodology, finding that subjects indicated they would be more likely than their peers to perform moral behaviors, but no more likely to perform intellectual behaviors. A theoretical framework is proposed in which people's positive beliefs about themselves are constrained by the publicity, specificity, and objectivity of the dimensions on which these beliefs are held.

Before being drafted and refusing induction into the army in 1967, heavyweight boxing champion Muhammad Ali was asked whether he had really failed the army mental examination or whether he had

This research was supported by grant BNS 83-02674 from the National Science Foundation to David Messick. Correspondence and requests for reprints should be sent to David M. Messick, Department of Psychology, University of California, Santa Barbara, CA 93106.

deliberately faked a poor performance in order to stay out of the service. Responding with a wit that belied his claim that the failure was genuine, Ali quipped, "I only said I was the greatest, not the smartest" (Ali, 1975, p. 129). The research reported in the present paper indicates that Ali's self-perception that he is better but not necessarily smarter than others is widely shared. Muhammad Ali was, and remains, a unique figure in many ways. However, he is not alone in his self-evaluation.

The present research begins with Messick, Bloom, Boldizar, and Samuelson's (1985) study of our perception that while both we and others perform fair behaviors more frequently than unfair behaviors, that difference is larger for ourselves than other people. It also follows a study by Goethals (1986) that found a similar result, namely, that people perceive that their own altruistic behaviors are performed less frequently by others than they actually are. Both sets of findings indicate that we think our own behavior is better, in the sense of being less often selfish, than that of others. The purpose of the present research was to find out if the perception that we are better than others extends to thinking that we are also smarter.

Messick *et al.* discovered a phenomenon that they labeled the *differential slope relationship*. This relationship may be described as follows: There is a positive relationship between the judged fairness of a behavior and the judged frequency with which the behavior occurs, and the strength of that relationship is stronger for judgments about ourselves than it is for judgments about others. Put somewhat differently, if one imagines a monotonic increasing function relating the mean frequency of an act to its mean fairness, the slope of that function would be steeper when the acts were self-ascribed than when they were ascribed to others. Liebrand, Messick, and Wolters (1986) replicated the Messick *et al.* pattern of results using Dutch-speaking subjects in the Netherlands.

Both of these studies focused only on the domain of fairness, however, a domain in which other kinds of cognitive biases have already been studied (Harris & Joyce, 1980; Messick & Sentis, 1983; Ross & Sicoly, 1979). The purpose of the present experiments was to determine whether the dual slope relationship generalizes to other dimensions.

We chose two different dimensions to explore in this study. First, we chose the general good–bad dimension, thinking that the fair–unfair results reported in our previous papers might reflect a more pervasive bias in the moral evaluation of ourselves and others. If the fair–unfair dimension is just one aspect of a more basic moral evaluation dimension, we would expect to replicate our previous pattern of results with the good–bad dimension.

For several reasons, we also decided to use the intelligent–unintelligent dimension. First, as Rosenberg and Sedlak (1972) have shown, it is a basic component of a more general evaluative dimension, which is distinct from moral evaluation. Second, as Reeder and Brewer (1979) have argued, a dimension that has an ability component has a different spectrum of attributional implications than one that lacks an ability component. When a person performs a behavior that requires ability, one can infer that the person possesses the ability. When a person performs a behavior such as being fair that does not require an ability component, it is less easy to make a corresponding inference. Thus a smart behavior requires intelligence, but a moral behavior does not necessarily require morality. This tighter correspondence between intelligent acts and inferences of intelligence permits less interpretational ambiguity than is possible with moral judgments. As a result, we expect to find a smaller self-serving bias with intelligence than with the moral dimension.

Our theoretical framework assumes that people prefer to hold positive rather than negative beliefs about themselves. This may mean being better than average on evaluative dimensions (Myers & Ridl, 1979) or having a greater than base-rate probability of performing positive acts (Goethals, 1986). This preference to believe that one is somewhat superior to one's peers, however, is tempered by two factors. First, not all dimensions are important. If one does not view a dimension as having evaluative meaning, one should not want to be distinctive on that dimension (Campbell, 1986). Beggan, Messick, and Allison (1988) demonstrated this phenomenon within the domain of cooperation and competition. Liebrand, Jansen, Rijken, and Suhre (1986) presented evidence that cooperators perceive the cooperative–competitive dimension as an evaluative dimension with the cooperative pole viewed as "good" and the competitive pole valued as "bad." Competitors, on the other hand, tend to view the dimension in terms of power. Thus, when subjects who had been classified as cooperators wrote lists of examples of cooperative and competitive acts, they wrote more first-person cooperative examples than first-person competitive ones. Competitors, in contrast, used first-person examples equally often for cooperative and competitive behaviors.

The second factor that constrains people's tendencies to exaggerate their positive qualities is the extent to which the quality is publicly and objectively verifiable. The less ambiguous, private, or subjective the attribute is, the less subject it is to self-serving exaggeration (Felson, 1981). Efforts to distort a specific, public, or objective position run the risk of detection by others and subsequent embarrassment. A deeper reason why people avoid self-serving distortions on dimensions where exaggerations are obvious is that people must believe

in their own veracity and objectivity in order to take their own opinions seriously (Gilbert & Cooper, 1985). In other words, people want their beliefs to have the force of truth, and that is difficult with beliefs that are continually contradicted by experience. It would be small gratification for a lunchtime jogger to pretend to be able to run a 4-minute mile or for a student of modest academic accomplishments to claim to be a National Merit Scholar. People might engage in such fantasies in Walter Mittyish episodes, but it is doubtful that such beliefs could become a stable part of an individual's self-image.

We propose that the dimension of intelligence is more specific, public, and objective than the dimension of moral goodness and that for that reason subjects should show smaller egocentric biases with the former than with the latter. The research that will be described consists of three studies. In the first, following the procedures of Messick *et al.* (1985), subjects will be asked to write lists of behaviors for both themselves and others that are either good or bad, or intelligent or unintelligent.

In the second study, samples of the written behaviors will be drawn and these samples will be rated for both goodness or intelligence and frequency. We will look for evidence of the differential slope relationship and see if that relationship is different for the good–bad versus intelligent–unintelligent behaviors. In the final study, following Goethals (1986), subjects will indicate the probability of themselves and a fellow student performing a desirable behavior. We will determine whether people think that they are more likely than others to perform desirable behaviors and whether they think so to an equal extent in both moral and intellectual domains.

STUDY 1

METHOD

Subjects. The subjects were 112 introductory psychology students from the University of California, Santa Barbara (UCSB) who participated to fulfill a course requirement. The subjects were run in groups ranging in size from 5 to 10 people.

Materials and Procedure. One half of the subjects were asked to write examples of good and bad behaviors, while the other half were to list intelligent and unintelligent behaviors. Each subject was given a questionnaire containing a cover sheet and two additional pages. For subjects in the good–bad condition, one of these two additional pages contained the following instructions:

> In the space below, please write as many things that you can think of that you do, or that other people do, that you would describe as good. If you think that you do these things more often than others, begin the sentence with "I". If you think that others do these things more often than you do, then start the sentence with "They". For example, "I work hard to achieve my goals," or "They help other people." You will be given five minutes for this task.

On the top of the other page, the following paragraph asking for bad behaviors was printed:

> In the space below, please write as many things that you can think of that you do, or that other people do, that you would describe as bad. If you think that you do these things more often than others, begin the sentence with "I". If you think that others do these things more often than you do, then start the sentence with "They". For example, "I steal from others," or "They are rude and insulting to others." You will be given five minutes for this task.

Subjects in the intelligent–unintelligent condition were given analogous instructions. In this case, the examples were "I can solve difficult problems" and "They organize their time wisely." The two examples of unintelligent actions were "I get involved with people who are bad for me" and "They spend money stupidly."

Half of the subjects were asked to write positive behaviors prior to generating negative behaviors. The order was reversed for the other half. When the subjects had answered all the questions on the cover sheet, the experimenter asked them to turn the page and read the instructions. He then timed them for 5 minutes, after which he told them to turn the next page. He then timed them for 5 more minutes. To minimize self-presentational tendencies, subjects were told that their answers would be anonymous, and they were not asked to write their names or other identifying information on the answer forms or cover sheet.

From the subjects' lists we recorded the number of positive and negative entries that began with "I" and the number that began with "They." In a few instances, subjects began a sentence with "We." In these cases we recorded one half for "I" and one half for "They."

RESULTS

The mean number of positive and negative behaviors that were listed in the first and third person are presented in Table 1. For each subject we computed the difference between the number of "I" and "They"

TABLE 1
Study 1: Mean Number of Good and Bad Behaviors and Intelligent and Unintelligent Behaviors in the First and Third Person for the Two Orders

	GOOD			BAD		
ORDER	I	THEY	TOTAL	I	THEY	TOTAL
Good–Bad	7.61	4.29	11.90	4.71	7.32	12.03
Bad–Good	5.79	4.04	9.83	3.54	6.68	10.22
Mean	6.70	4.17	10.87	4.13	7.00	11.13
	INTELLIGENT			UNINTELLIGENT		
ORDER	I	THEY	TOTAL	I	THEY	TOTAL
Int–Unint	5.07	3.93	9.00	3.39	5.53	8.92
Unint–Int	5.85	4.07	9.92	4.17	5.00	9.17
Mean	5.46	4.00	9.46	3.78	5.27	9.05

statements for both good and bad behaviors and intelligent–unintelligent ones. The I–They differences were first analyzed in a separate 2 × 2 (good–bad by order) ANOVA. The analysis revealed a large main effect for good versus bad, F (1, 54) = 77.62, $p < .001$. On the average, subjects generated 2.54 more first- than third-person good behaviors, but 2.88 fewer first- than third-person bad behaviors. The same ANOVA was run for intelligent and unintelligent behaviors, and it uncovered the same effect: Subjects wrote 1.46 more first- than third-person intelligent behaviors, but 1.48 fewer first- than third-person unintelligent behaviors, F (1, 54) = 17.71, $p < .001$.

To examine possible differences between the good and bad behaviors and the intelligent and unintelligent behaviors, we combined the two data sets and computed a 2 (type of behavior) × 2 (order) × 2 (pole) ANOVA with repeated measures on the last factor. The dependent variable was the difference between the frequency of "I" and "They" behaviors. The analysis revealed three significant effects. First, there was the expected main effect for pole, F (1, 108) = 80.52, $p < .001$, indicating that subjects generated more "I" than "They" positive behaviors and more "They" than "I" negative behaviors. The analysis also yielded an interaction between type of behavior and order, F (1, 108) = 5.45, $p < .022$. This interaction indicates that subjects in the good–bad condition wrote an average of .36 more "I" than "They" behaviors when positive behaviors preceded negative ones but wrote .70 more "They" than "I" behaviors when the order was reversed. Conversely, subjects asked to generate intelligent and unintelligent

behaviors wrote .50 more "They" than "I" behaviors when positive behaviors preceded negative ones but wrote .48 more "I" than "They" behaviors when the order was reversed.

The third significant effect was a pole × type of behavior interaction, F (1, 108) = 7.00, $p < .010$. This effect reveals that the difference between the number of "I" and "They" behaviors generated was not as pronounced in the intelligent–unintelligent condition as in the good–bad condition. That is, while subjects wrote more first- than third-person intelligent behaviors and fewer first- than third-person unintelligent behaviors, these differences were smaller than those in the good–bad condition.

Further analysis of the data in Table 1 revealed one unexpected effect, an interaction between type of behavior, order, and pole, F (1, 54) = 5.68, $p < .021$. While more first- than third-person good items and more third- than first-person bad items were generated, the absolute magnitude of both of these differences was smaller in the second than in the first position. However, this pattern is reversed in the intelligent–unintelligent condition: While more first- than third-person intelligent behaviors and more third- than first-person unintelligent behaviors were listed, the absolute magnitude of the difference was smaller in the first than in the second position.

DISCUSSION

The basic egocentric bias reported by Messick *et al.* (1985) and Liebrand *et al.* (1986b) was found for both the good–bad and the intelligent–unintelligent dimensions. The subjects wrote more first-person than third-person positive behaviors and more third-person than first-person negative behaviors. The pole × type of behavior interaction indicates that the magnitude of the bias was smaller for the intelligent–unintelligent dimension than for the good–bad one.

Both Messick *et al.* (1985) and Liebrand *et al.* (1986b) found an interaction between the magnitude of the bias and order. In both cases, the egocentric bias was more pronounced in the first position than in the second position. The absolute values of difference between the "I" and "They" frequencies was greater when the lists of fair and unfair acts were written first rather than second. Exactly the same pattern was found for the good–bad dimension in the present study. The absolute values of the differences between the "I" and "They" frequencies were 3.32 and 3.14 for good and bad behaviors in the first position, but only 1.75 and 2.61 for good and bad behaviors in the second position.

For the intelligent–unintelligent dimension, the pattern is the opposite. In the first position, the absolute value of the difference in "I" versus "They" frequencies was 1.14 and .83 for intelligent and unintelligent behaviors respectively, but in the second position these differences were 1.78 and 2.14 respectively.

This difference between the two dimensions may be an important one. We explain why in the general discussion.

STUDY 2

METHOD

Subject. The subjects were 122 UCSB introductory psychology students who participated to fulfill a course requirement. Subjects were run in groups ranging in size from 5 to 20 persons.

Stimulus Materials and Procedures. Subjects were given a questionnaire that contained a cover sheet for recording biographical information. On the second page subjects were given a set of instructions that varied in terms of the type of behaviors to be rated (either good–bad or intelligent–unintelligent) and in the type of rating (either frequency or the degree of goodness or intelligence). The instructions for frequency ratings were as follows:

> On the following pages we have listed a number of behaviors that other students wrote as being good (intelligent) or bad (unintelligent). We are interested in your impression of how often these behaviors occur. We would like you to rate each of the following behaviors in terms of how frequent or infrequent you think it is. We would like you to use the following scale:
>
> 1 = very frequent
> 2 = rather frequent
> 3 = rather infrequent
> 4 = very infrequent
>
> On the sheets that follow, please write the number that best reflects your attitude about the behavior. Write the number of the left of each of the items. Please rate each behavior and please try to be as accurate as possible. If you have any questions, please raise your hand.

The instructions for indicating the degree of goodness or intelligence of behaviors were worded as follows:

> On the following pages we have listed a number of behaviors that other students wrote as being good (intelligent) or bad (unintelligent).

We are interested in your impressions of these behaviors. We would like you to rate each of the following in terms of how good (intelligent) or bad (unintelligent) you think it is. We would like you to use the following scale:

1 = very good (intelligent)
2 = somewhat good (intelligent)
3 = slightly good (intelligent)
4 = neither good (intelligent) nor bad (unintelligent)
5 = slightly bad (unintelligent)
6 = somewhat bad (unintelligent)
7 = very bad (unintelligent)

The remainder of the instructions were identical to the frequency instructions. There were approximately equal numbers of subjects making each of the four ratings.

The behaviors that subjects rated were randomly selected from the pool of 1200 behaviors generated in the first study. There were 20 behaviors in each of eight distinct types. The statements were either good, bad, intelligent, or unintelligent behaviors, and they began either with "I" or "They." Subjects rated either the 80 good and bad acts, or the 80 intelligent and unintelligent acts. Roughly, one fourth of the subjects rated the goodness of the 80 good and bad behaviors and one fourth rated their frequency. One fourth rated the intelligence of the 80 intelligent and unintelligent behaviors and one fourth rated their frequency. The statements were randomly ordered, 20 to a page, and the pages were randomly ordered within subjects' booklets. Interviews with subjects in our previous studies indicated that subjects treated the "I" statements as self-referential, so that when rating the frequency of a behavior like, "I gossip a little," they were referring to their own behavior, not to the frequency with which other subjects wrote "I gossip a little."

RESULTS

The mean goodness and intelligence ratings are provided in Table 2, and the mean frequency ratings are shown in Table 3.

Goodness and Intelligence Ratings. For each subject we computed the mean rating of each of the four sets of 20 behaviors rated (I-good, They-good, I-bad, They-bad, or I-intelligent, They-intelligent, I-unintelligent, They-unintelligent). These judgments, shown in Table 2, were then analyzed in a 2 (type of behavior) × 2 (person) × 2 (pole) ANOVA, with repeated measures on the last 2 factors. The ANOVA uncovered several significant effects. First, there was a main effect for

TABLE 2
Study 2: Mean Judgments of Goodness or Intelligence of Behaviors That Were Written as Examples of Good, Bad, Intelligent, or Unintelligent Behaviors Associated with Self or Others

	TYPE OF BEHAVIOR		
	GOOD–BAD	INT–UNINT	MEAN
I-positive	1.71	2.62	2.16
I-negative	5.31	4.72	5.02
They-positive	1.82	2.46	2.14
They-negative	5.91	5.19	5.56
Mean	3.69	3.75	3.72

person, F (1, 59) = 74.15, $p < .001$, indicating that "I" behaviors were rated as more positive (M = 3.59) than "They" behaviors (M = 3.84). Furthermore, there was an interaction between type of behavior and person, F (1, 59) = 11.38, $p < .002$. This effect shows that the difference between the ratings of first- (M = 3.51) and third-person (M = 3.87) good and bad behaviors was greater than the difference between first- (M = 3.67) and third-person (M = 3.83) intelligent and unintelligent behaviors.

The analysis expectedly revealed a main effect for pole, F (1, 59) = 489.71, $p < .001$, indicating that the positive behaviors written in Study 1 were rated more positively than the negative behaviors. However, we also found an interaction between pole and type of behavior, F (1, 59) = 25.54, $p < .001$. Subjects were less extreme in their judgments of intelligent (M = 2.54) and unintelligent (M = 4.96) behaviors than in their judgments of good (M = 1.77) and bad (M = 5.61) behaviors.

TABLE 3
Study 2: Mean Frequency Ratings of Behaviors That Were Written as Examples of Good, Bad, Intelligent, or Unintelligent Behaviors Associated with Self and Others

	TYPE OF BEHAVIOR		
	GOOD–BAD	INT–UNINT	MEAN
I-positive	2.09	2.43	2.26
I-negative	2.64	2.57	2.61
They-positive	2.47	2.55	2.51
They-negative	2.59	2.68	2.63
Mean	2.45	2.56	2.50

There was also an interaction between pole and person, F (1, 59) = 76.88, $p < .001$. While "I" and "They" positive behaviors were seen as equally positive (M = 2.16 and 2.14, respectively), "They" negative behaviors were rated more negatively (M = 5.56) than "I" negative behaviors (M = 5.02).

Frequency Ratings. Subjects' frequency ratings, analyzed in a 2 (type of behavior) × 2 (person) × 2 (pole) ANOVA, yielded main effects for all three factors. Good and bad behaviors were rated as more frequent (M = 2.45) than intelligent and unintelligent behaviors (M = 2.56), F (1, 59) = 6.88, $p < .012$. First-person behaviors were judged as occurring more frequently (M = 2.43) than third-person behaviors (M = 2.57), F (1, 59) = 23.86, $p < .001$, and positive behaviors were rated as more frequent (M = 2.39) than negative ones (M = 2.62), F (1, 59) = 10.28, $p < .003$.

The analysis also uncovered two interactions. First, there was an interaction between person and pole, F (1, 59) = 7.52, $p < .009$. This effect indicates that while "I" and "They" negative behaviors were judged as equally frequent, "I" positive behaviors were rated as more frequent than "They" positive behaviors. However, this interaction is qualified by a three-way interaction between type of behavior, person, and pole, F (1, 59) = 6.64, $p < .013$. As Table 3 shows, this finding reflects a sizable difference in the positive and negative "I" frequency ratings and a much smaller difference for the "They" ratings (the differential slope pattern). These differences for the intelligence dimensions were virtually identical. A simple effects test confirmed this interpretation: The two-way interaction between person and pole was significant for good–bad behaviors ($p < .01$), but not for intelligent–unintelligent behaviors ($F < 1$).

DISCUSSION

The positivity ratings—goodness or intelligence—replicate the fairness results of Messick *et al.* (1985) and Liebrand *et al.* (1986b), with the effects for the intelligence dimension being somewhat weaker than those for goodness. The interaction between type of behavior and pole shows that the difference in evaluation between the positive and negative poles is greater for good–bad than for intelligent–unintelligent. Moreover, the type by person interaction shows that the difference between "I" and "They" behaviors is also greater for good–bad acts than for intelligent–unintelligent ones.

It is with the frequency ratings that we seek evidence for the differential slope relationship. An inspection of Table 3 makes the meaning

of the crucial three-way interaction very clear. For the intelligent–unintelligent dimensions, the difference in the rated frequency of the I-positive and I-negative behaviors (M = .14) is virtually identical to that for the "They" behaviors (M = .13). Thus for this dimension the slope difference is essentially zero.

With the good–bad dimension, however, we find the differential slope pattern. The difference between the I-good and I-bad frequencies (M = .55) is significantly larger than that between the They-good and They-bad acts (M = .12). Thus for the good–bad dimension we find the differential slope pattern, but not for the intelligent–unintelligent dimension.

STUDY 3

The first two studies indicate that people believe that they engage in positive behaviors more frequently than others and that this perception is stronger in the general domain of goodness than in the domain of intelligence. A series of studies reported by Goethals (1986) presents findings that are similar to those of Messick *et al.* (1985). Goethals found that people underestimated the percentage of their peers who would perform the moral or altruistic acts that they themselves would perform (*e.g.*, helping a friend, giving blood, acting cooperatively) and overestimated the percentage of their peers who would perform the more self-centered, egocentric, or selfish acts (*e.g.*, refusing to help a friend, refusing to give blood, and acting competitively). The overestimation of the frequency of one's moral behavior is similar to the findings above, but produced by a notably different methodology. Study 3 investigates whether the tendency to underestimate the frequency of others performing desirable acts is greater for moral than intelligent acts.

METHOD

Subjects. The subjects were 62 introductory psychology students at UCSB, whose participation fulfilled a course requirement. The subjects were run in groups of four to six persons each.

Design. The experimental design was a 2 (Type of Behavior: Moral vs. Intellectual) × 2 (Target: Self vs. Average UCSB Student) × 2 (Order of Judgment: Self then Average UCSB Student vs. Average UCSB Student then self). The first two factors were within-subject factors, while the last was a between-subjects factor.

Stimulus Materials. Each subject read eight one-paragraph vignettes. Each vignette described a fictitious person who was confronted with either a moral or an intellectual problem. Four of the eight vignettes described a moral problem, while four contained an intellectual problem.

The four moral vignettes featured (a) Alice, who is driving to an important job interview and must choose between helping a stranded elderly couple or arriving at the interview on time; (b) Charlie, who is sharing a pizza with a friend and must decide whether to eat the larger or the smaller of two remaining pieces while his friend is momentarily away from the table; (c) Judy, who must choose between studying for an exam and driving a friend who has been called home for an emergency to the airport; and (d) Mark, who is coaching a Youth Center basketball team in the championship game and must decide whether to play the five best players the entire game or allow every player on the whole team to play.

The four intellectual vignettes featured (a) Bill, who is on a Trivial Pursuit team and must answer at least 60% of the questions correctly if he is to help his team; (b) Debbie, who hears on the radio that she can win a free color T.V. set if she correctly completes the Sunday Los Angeles Times crossword puzzle within 3 hours; (c) Sheila, who is applying for a job as a law clerk and must correctly define at least 60% of the words in a vocabulary test; and (d) Steve, who is driving to an unfamiliar section of New York City when he realizes that he forgot his directions and must rely on his memory of those directions if he is to arrive at his destination.

Procedure. When subjects arrived, they were given the eight vignettes in alphabetical order, i.e., Alice, Bill, Charlie, Debbie, Judy, Mark, Sheila, and Steve. Each vignette was on a different page. Subjects were told to read each vignette and to answer the question at the bottom of each page.

For one half of the subjects, the question at the bottom of each page asked subjects to estimate the chances that they would perform the moral or intellectual behavior in question in the vignette. For example, for the Alice vignette, subjects were asked, "If you were in Alice's position, what are the chances, from 0 to 100%, that you would stop and help the stranded elderly couple?" Similarly, in the Bill vignette, subjects were asked, "If you were in Bill's position, what are the chances, from 0 to 100%, that you would be able to answer 60% of the questions correctly?" After subjects answered this question for each vignette, the experimenter asked them to go back to each vignette and estimate the chances that the average UCSB student would perform the behavior. Subjects were asked to write their answers to this second question on the same page, below their responses to the first question.

For the remaining half of the subjects, they first estimated the chances that the average UCSB student would perform the behavior and then judged the likelihood that they themselves would perform the behavior.

RESULTS

Subjects' mean percentage estimates for themselves and for the average UCSB student are displayed in Table 4. These percentages were subjected to a 2 (Order) × 2 (Type of Behavior) × 2 (Target) ANOVA with repeated measures on the last two factors. The analysis revealed three significant effects. First, there was a main effect for Target, $F(1, 60) = 13.41$, $p < .001$, indicating that, overall, subjects believed that they would be more likely to perform the given behavior ($M = 55.13$) than would the typical UCSB student ($M = 50.44$). Second, there was an interaction between Order and Type of Behavior, $F(1, 60) = 5.45$, $p < .03$. The order in which subjects were asked to make their estimates appeared to affect these estimates more for moral be-

TABLE 4
Study 3: Mean Percentage Estimates as a Function of Order, Type of Behavior, and Target of the Estimates

	ORDER			
	SELF–UCSB		UCSB–SELF	
MORAL	SELF	UCSB	SELF	UCSB
Alice	53.21	40.47	34.83	31.97
Charlie	74.07	48.24	62.30	43.91
Judy	67.57	51.06	59.26	58.09
Mark	68.98	50.82	63.26	50.09
Mean	65.96	47.65	54.91	46.02
	ORDER			
	SELF–UCSB		UCSB–SELF	
INTELLECTUAL	SELF	UCSB	SELF	UCSB
Bill	32.83	47.22	42.78	47.50
Debbie	35.92	41.87	32.43	44.75
Sheila	60.71	62.72	72.35	69.71
Steve	61.92	53.84	59.64	58.36
Mean	47.85	51.41	51.80	55.08

haviors than for intellectual behaviors. Specifically, when the behaviors were moral ones the differences between self and average UCSB student percentages were greater when the order was self–UCSB student than when the order was the reverse.

Of greatest interest, however, was the interaction that emerged between Type of Behavior and Target, F (1, 60) = 34.38, $p < .001$. This interaction qualifies the Target main effect in that it clearly demonstrates that subjects believed that they would be more likely to perform a moral behavior than would the average UCSB student, but would be slightly less likely to perform an intellectual behavior.

DISCUSSION

Again, we find that people respond differently to the moral dimension than to the intellectual dimension. With the moral scenarios, people exhibit the so-called *illusion of uniqueness*, the tendency to think that they are more likely than their peers to perform a somewhat costly moral act. This illusion does not exist with the four intellectual acts that we used, however. Our subjects did not indicate that they were more likely than their peers to be able to perform an act requiring intellectual skill or ability.

GENERAL DISCUSSION

Our interpretation of these basic findings is that the weaker biases found in the intelligence domain are attributable to the greater publicity, specificity, and/or objectivity of behaviors signaling intelligence than of behaviors indicating morality. To set the stage for the interpretation of some of the unexpected findings that emerged from these studies, we will present a brief overview of our theoretical perspective.

People wish to hold positive beliefs about themselves. These beliefs are often that they are at least average on important dimensions and possibly above average. In order for people to take their positive self-evaluations seriously, however, these evaluations must be credible, as Gilbert and Cooper (1985) have pointed out. This is to say that people cannot simply make up fantasies about themselves that would be blatantly contradicted by external facts and events. People have to believe that they are objective with regard to the beliefs that they hold about themselves.

The need to believe that self-evaluations are objective places important constraints on them, and this fact highlights the need to

discover the occasions on which biased self-evaluations could be observed *but are not*. For instance, Messick *et al.* (1985) found that subjects did not alter the rated fairness of behaviors as a function of whether the behavior was in the first person, for example, "I cheat on tests," or in the third person, "They cheat on tests." The rated frequency of the behavior changed markedly with the pronoun, however. A full theory of self-serving biases will have to explain not only why we believe we do unfair things less often than others, but also why the rated fairness of the behavior itself does not change. Our concept of objectively constrained biases does precisely this. The three central elements of the theory follow.

First, people do not exaggerate their positions on all dimensions. As Beggan *et al.* (1988) showed, people who do not associate an evaluative meaning with a dimension are unlikely to care about their position on the dimension. The results of the present experiment support the hypothesis that people do not exaggerate their positions on dimensions that are public, specific, and more or less objective. One consequence of not exaggerating one's position on all dimensions is that the self-evaluation may be perceived as objective because it manifests discrimination. A self-image in which one was better than others on all dimensions would not be a credible one. A self-image in which one is not better than others on some dimensions, those that one does not care about, for instance, or that are highly objective, is much more credible. Thus the lack of bias on some dimensions provides positive evidence that the self-evaluation is discriminating and objective.

Second, to maintain the belief that a self-image is credible, people must present themselves with direct evidence of its objectivity. In the paragraph above, we indicated one source of such evidence. A second source may derive from the failure of subjects to alter their evaluative ratings of behavior as a function of whether they pertain to the self or others, while markedly changing their frequency ratings. The contrast between self and others in this case appears to have a counterfactual form. It would be as bad for me to cheat on tests as it is for them to cheat on tests, but I don't do it, at least not as often as they do. Acknowledging that the quality of the behavior would be the same for self and others constitutes evidence of objectivity; the differentiation is done along the frequency dimension.

We also present direct evidence of our objectivity when we admit that we do negative things and acknowledge that others do positive things. An unconstrained self-serving bias would predict that when writing good (intelligent) behaviors only first-person examples should be observed, and that when writing bad (unintelligent) behaviors only third-person acts should appear. As Messick (1987) has noted, these

written behaviors consistently occur around 60% of the time, not 100%. Messick has proposed a mechanism based on the idea that people need to maintain an objective self-image that predicts frequencies of 61.8% first-person positive and third-person negative behavior.

The third factor that supports the illusion of objective self-images is that some of the processes that produce bias occur out of conscious awareness and do not, therefore, challenge the validity of the self-image (Gilbert & Cooper, 1985). Liebrand *et al.* (1986b) found, for instance, that subjects tended to remember a higher proportion of unfair behaviors that began with "They" than either unfair behaviors that began with "I" or fair behaviors. The greater memorial availability of "They-unfair" behaviors is not likely to result from a conscious strategy, but its consequence is the association of others with negative behaviors.

People have much more access to their own subjective states than to those of others. People may, therefore, evaluate themselves on their internal states to a greater degree than they do others. In the present study, we tested this idea by examining the sample of good, bad, intelligent, and unintelligent behaviors for omissions and intents. The first of these, omissions, were behaviors, written as positive, either good or intelligent, that were failures to do negative things. Examples of these are, "I do not cheat" and "I don't smoke." People are, we hypothesized, more likely to credit themselves for resisting the temptation to do a negative behavior than they are to credit others, primarily because the nonbehaviors of others have virtually no perceptual salience. Of the 40 positive "I" behaviors sampled in Study 2, 7 were omissions. None of the 40 "They" positive behaviors was an omission.

People also credit themselves for their good intentions, but they cannot credit others easily because others' intentions are less conspicuous. Three of the 40 positive "I" behaviors were intents, for example, "I try to be kind to others," whereas none of the "They" positive behaviors were intents.

It is highly unlikely that our subjects were consciously aware that they were writing omissions or intents as instances of positive things for themselves but not for others. In the same vein, we suspect that subjects were unaware of the fact that they tended to use the frequency dimension rather than the evaluative dimension to differentiate themselves from others. This tendency also partly explains why "I" negative behaviors were rated more positively than "They" negative behaviors in Study 2. Many of the behaviors that were written included frequency-related adverbial modifiers like "too much," "often," or "not enough." If these adverbs were used differently in describing "I" and "They" behaviors, different evaluative ratings could result. Two examples will

serve to illustrate the point. One of the "I" unintelligent behaviors that we sampled was, "I don't read enough." Its intelligence rating was 4.73. One of the "They" unintelligent behaviors was, "They do not read." Its rating was 5.57. The inclusion of the word "enough" in the first person act increases the evaluative rating nearly three quarters of a scale unit. Likewise, the difference between "I gossip a little" and "They gossip too much" is the difference between intelligence ratings of 4.17 and 5.00. The difference has to do with the frequency, not the quality, of the act.

One reason why the frequency dimension appears to be the major dimension of differentiation of self from others is that the meanings of frequency and probability expressions are generally vague and idiosyncratic (*e.g.*, Budescu & Wallsten, 1985). Words like "sometimes" or "a little" can refer to a wide latitude of frequencies, and there is great interjudge variability in the frequencies that words like these evoke. It is tempting to propose that the virtue of the frequency dimension and its associated language for differentiating the self from others derives from its very vagueness. The vagueness and lack of shared meaning of frequency and probability terms, qualities that interfere with precise communication, are exactly those qualities that insulate a statement or belief against contradiction.

There are other theoretical perspectives that are relevant to the results we have presented. Most notably, Alicke (1985) has shown that the rated applicability of trait adjectives to self and average college students depends not only on the desirability of the adjectives but also on their judged controllability. Alicke's hypothesis was that controllable traits are more personally revealing than uncontrollable traits. Therefore, it is more important to be above average on desirable and below average on undesirable controllable traits than on uncontrollable ones. Said differently, the self minus other ratings of applicability should change more as a function of desirability when the traits are controllable than when they are not. If morality were construed as controllable and intelligence as uncontrollable, Alicke's (1985) hypothesis would predict the same pattern of judgments that we did. An inspection of the traits that Alicke used in his study is suggestive in that the traits "intelligent" and "intellectual" were both highly desirable, low-controllable traits.

Campbell (1986) has argued that while people want to be distinctive with regard to abilities, they want to be less so in the domain of opinions. A similar factor might be at work in our data. However, an important aspect of the opinion–ability distinction is that abilities have a universally shared evaluation, in the sense that more of an ability is better than less of one, whereas no such universality exists in the

evaluation of opinions (Sherman, Chassin, Presson, & Agostinelli, 1984). In our data, evaluation in both the moral and intellectual domains appears to have been more universal than not in the sense that subjects clearly differentiated the positive and negative behaviors in their ratings. Thus Campbell's distinction between opinions and abilities appears less pertinent than Alicke's notion of controllability to our morality–intelligence distinction. The final word on the relative merits of these approaches, compared to the one offered here, will have to await further research.

We will conclude by offering explanations for the order interactions that we found in Study 1 and Study 3. In Study 1, we found that the egocentric bias was stronger in the first than in the second position when subjects wrote lists of good and bad acts, but that the effect was reversed when they wrote lists of intelligent and unintelligent behaviors. Both Messick *et al.* (1985) and Liebrand *et al.* (1986b) reported order effects identical to those found in the current study with good–bad behaviors. One interpretation of this effect is that the desire to differentiate self from others in the second position is reduced because that differentiation has already been made in writing the first list. This hypothesis implies that the need to view the self as better than others can be at least partly sated, and it also suggests that doing so, at least in this writing task, is effortful. The order effect does not appear to be a general fatigue effect because we do not see a reduction in the total number of behaviors listed in the second as opposed to the first position. It is the structuring, not the number of acts, that changes.

The reversal of the order effect for the intelligent–unintelligent task was unexpected. The fact that there was a larger differentiation between self and others in the second than in the first position may indicate that rehearsal or practice is required when the evaluative dimension is relatively objective. The constraints imposed by objectivity may be stronger for the first list than for the second. Taken together with the hypothesis offered in the preceding paragraph, the idea emerges that when subjects do the writing task for the first time, either the tendency to differentiate or the tendency to display objectivity predominates, depending on which of these two antagonistic requirements is evoked by the task: differentiation by good–bad, objectivity by intelligent–unintelligent. The relative emphasis on one of the two processes diminishes with the lists written in the second position. Therefore the second position lists from the two domains should be more similar in composition than the lists written in the first position. This implication is clearly borne out by the data. The proportion of first-person positive and third-person negative behaviors written in

the first position is .646 for good–bad and .554 for intelligent–unintelligent. These same proportions for the second-position items are .600 and .605, respectively. There is a large difference in the first position and virtually none in the second.

The order effect that was observed in Study 3 appears to be somewhat different from that in Study 1, which is not surprising since the tasks are totally different. Inspection of Table 4 reveals that the major effect of order was on the estimates of the probability that self would perform the moral behaviors. When subjects must make a judgment of their own likelihood of performing one of these moral acts after having made estimates for UCSB students in general, the latter judgment may serve as an anchor for the judgment about self. To display a superiority over UCSB students, a subject would only need to exceed this anchor by a small amount. No such context is available to constrain the judgment about self when that judgment comes first. The fact that initial self-judgments are not higher when made about behaviors in the intellectual domain suggests that other constraints are already operating, namely, those implied by the specificity, publicity, and objectivity of the intellectual domain.

To summarize, in order to understand self-serving biases, it is necessary to understand when they are weak or nonexistent. In other words, we need to understand how it is that, like Muhammad Ali, we tend to think that we are better, albeit not necessarily smarter, than others.

REFERENCES

Ali, M. (1975). *The greatest: My own story*. New York: Random House.

Alicke, M. E. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49*, 1621–1630.

Beggan, J. K., Messick, D. M., & Allison, S. T. (1988). Social values and egocentric bias: Two tests of the Might over morality hypothesis. *Journal of Personality and Social Psychology, 55*, 606–611.

Budescu, D. V., & Wallsten, T. S. (1985). Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes, 36*, 391–405.

Campbell, J. D. (1986). Similarity and uniqueness: The effects of attribute type, relevance, and individual differences in self-esteem and depression. *Journal of Personality and Social Behavior, 50*, 281–294.

Felson, R. B. (1981). Ambiguity and bias in the self-concept. *Social Psychology Quarterly, 44*, 64–69.

Gilbert, D. T., & Cooper, J. (1985). Social psychological strategies of self-deception. In M. Martin (Ed.), *Self-deception and self-understanding*. Lawrence, KS: University of Kansas Press.

Goethals, G. R. (1986). Fabricating and ignoring social reality: Self-serving estimates of consensus. In J. M. Olson, C. P. Herman, & M. P. Zanna (Eds.), *Social comparison and relative deprivation: The Ontario symposium* (Vol. 4). Hillsdale, NJ: Erlbaum.

Harris, R. J., & Joyce, M. A. (1980). What's fair? It depends on how you phrase the question. *Journal of Personality and Social Psychology, 38*, 165–179.

Liebrand, W. B. G., Jansen, R., Rijken, V., & Suhre, C. (1986a). Might over morality: Social values and the perception of other players in experimental games. *Journal of Experimental Social Psychology, 22*, 203–215.

Liebrand, W. B. G., Messick, D. M., & Wolters, F. J. M. (1986b). Why we are fairer than others: A cross-cultural replication and extension. *Journal of Experimental Social Psychology, 22*, 590–604.

Messick, D. M. (1987). Egocentric biases and the golden section. *Journal of Social and Biological Structures, 10*, 241–247.

Messick, D. M., Bloom, S., Boldizar, J. P., & Samuelson, C. D. (1985). Why we are fairer than others. *Journal of Experimental Social Psychology, 21*, 480–500.

Messick, D. M., & Sentis, K. P. (1983). Fairness and preference. *Journal of Experimental Social Psychology, 15*, 418–434.

Myers, D. G., & Ridl, J. (1979). Can we all be better than average? *Psychology Today, 13*, 89–92.

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review, 86*, 61–79.

Rosenberg, S., & Sedlak, A. (1972). Structural representations of implicit personality theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 235–297), New York: Academic Press.

Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology, 37*, 322–336.

Sherman, S. J., Chassin, L., Presson, C. C., & Agostinelli, G. (1984). The role of the evaluation and similarity principles in the false consensus effect. *Journal of Personality and Social Psychology, 47*, 1244–1262.