Status and Confidence, in the Lab *

Jeffrey V. Butler

Einaudi Institute for Economics and Finance, Rome, Italy

jeff.butler@eief.it

Abstract: It is widely recognized that confidence can have important economic consequences. While most of the focus has been on overconfidence, systematic variation in confidence can imply systematic variation in economic outcomes. Intriguingly, sociological and social psychological research suggests that being on the wrong side of inequality undermines confidence. This paper examines the link between inequality and confidence in a controlled, incentive-compatible laboratory setting. Inequality was introduced into an otherwise-standard social learning experiment by randomly assigning half of the subjects into a "high-paying" group and the other half into a "low-paying" group. The pay differences were lump-sum, making the economic incentives across groups identical. I found a significant link between inequality and confidence: being on the positive side of inequality increased confidence, while being on the negative side of inequality decreased confidence, relative to control sessions without inequality. This inequality-confidence link provides an alternative explanation for the type of self-reinforcing inequality present in, for example, U.S. educational attainment data—but which is typically attributed to external discrimination. It can also explain some patterns in college attainment that are inconsistent with an external discrimination story.

*I owe a great debt of gratitude to George A. Akerlof, Stefano DellaVigna and Shachar Kariv and Luigi Guiso for helpful comments. I am also grateful to to the National Science Foundation under Research Grant SES 04-17871 (Akerlof) and UC Berkeley's X-lab for invaluable financial support. I am also grateful to seminar participants at the European University Insitute and LUISS for comments and suggestions.

1 Introduction

Confidence, defined as a sanguine assessment of one's relative ability, is crucial for the myriad economically significant decisions in which rewards depend on relative performance criteria.[2] The class of decisions that depend on confidence runs the gamut from choosing whether to apply to college, to choosing one's career, to deciding whether and when to seek promotions and wage increases at various points along one's chosen career path.

Promotions and wage increases are, for instance, often explicitly based on relative (past-) performance criteria; and relative performance criteria such as SAT scores oftentimes factor heavily into college admissions decisions.[3] Expectations about future relative performance affect today's choices about whether to make the up-front and irretrievable investments of money, effort and time required for many of the most lucrative career paths such as becoming an entrepreneur.

Given the importance of confidence it is surprising that, despite a lively literature investigating the consequences of systematic *over*confidence (see, among others, Roll, 1986; Camerer and Lovallo, 1999; Malmendier and Tate, 2005,2008;), there is relatively little economic research addressing systematic patterns in the distribution of confidence.[4] Because decisions requiring confidence are ubiquitous, consequential and recur over the life cycle, systematic heterogeneity in confidence can imply significant long-term consequences—both for individuals and society. Repeatedly shying away from competitive situations can hurt individuals directly by impeding upward mobility. To the extent that individuals' lack of confidence is at odds with true relative ability, individual effects of the lack of confidence can impact the economy through inefficiently low levels of human capital attainment

[2]Provided, of course, that performance depends at least partially on ability. Given this, current confidence affects individuals' expectations of future relative performance, and, before true relative performance is revealed, confidence affects individuals' beliefs about relative current- and past-performance.

[3]More conventionally, because a large portion of the returns to college are conferred only upon degree completion—the so-called sheepskin effect—and degree completion is ultimately based on relative performance criteria (grade point averages), the decision to incur the substantial costs—in terms of money, effort and foregone wages—of attending college depends on expected future relative performance.

[4]Exceptions include several papers demonstrating that confidence varies with task difficulty (see, e.g., Moore and Cain, 2007); and a handful of papers documenting cross-cultural and gender-based variation in overconfidence (e.g., Niederle and Vesterlund, 2007 or Bengsston, Person and Willenhag, 2005).

and entrepreneurial activity.[5] On the other hand, persistently thinking too much of one's abilities—unwarranted *over*confidence—can hurt individuals and society through the kind of excessive risk-taking that is currently playing havoc with the world economy, and which plays out in the small every day in the form of high failure rates of small businesses (see, e.g., Camerer and Lovallo, 1999).

As a point of departure, there is considerable evidence from sociology and social psychology of what could be called a wealth effect with respect to confidence, in that individuals who presently "...differ in resources or rewards tend to create expectations for one another's competence and performance based on the basis of those rewards..." (Ridgeway, *et al*, 1998). The idea that people infer backwards from outcomes to ability in a way that justifies an unequal *status quo* has a long pedigree, spanning multiple social sciences, dating back to at least Allport (1954). Broadly, the phenomenon is explained by the combination of a basic human desire for justice, together with an aversion to cognitive dissonance, that requires individuals to rationalize existing social structures and outcomes as fair and legitimate.[6] In the current context the phenomenon implies the hypothesis that individuals who find themselves on the wrong side of inequality—on the lower rungs of the socio-economic ladder—will believe that their low position is due to their relative lack of ability. And, on the other hand, those at the top of the heap will believe they are there because of their relatively superior ability. In the real world, socioeconomic status is often taken to be a proxy for underlying inequality; in the remainder of the paper, therefore, for I will use the terms status and inequality interchangeably—despite the various other connotations of the word status. The phenomenon I will investigate can thus be restated simply as a positive relationship between status and confidence.

The economic importance of this status-confidence link is that it provides an explana-

[5] A secondary effect of the lack of confidence is to reduce the learning opportunities individuals have about their true relative abilities, adding an extra layer of self-reinforcement to a lack of confidence that is not present with *over*confidence.

[6] For a discussion of cognitive dissonance as well as other economically-relevant implications, see, e.g., Akerlof and Dickens, 1982. On backwards inference and its connection to a desire for justice, there are many sources. In social psychology, this tendency is called the Just World Hypothesis. A relatively recent survey can be found in Furnham, 2003. In sociology, the phenomenon is sometimes called System Justification Theory; for an overview and supporting evidence, see Jost, *et al*, (2002, 2003).

tion for the type of self-reinforcing economic underperformance that has long been a puzzle to economists—an explanation which does not rely directly on expectations of *extrinsic* discrimination, nor on differences in stable, underlying traits, nor on long histories of maltreatment. Low status individuals believe they are relatively less able *qua* low status, and are consequently less likely to embark on paths involving competition with higher-status counterparts. But these competitive paths are precisely the paths up and out of initial low status positions. The implication is that initial inequality persists naturally, rather than being one of many possible equilibria arising from, e.g., employers' initial discriminatory beliefs (Arrow, 1972); and apart from requiring a preference for discrimination (Becker, 1971.).

As one obvious example, consider the puzzle of persistent under-investment in human capital among the underclasses. This is commonly attributed to expectations of labor market discrimination based on race, for instance. However, if racial discrimination in the labor market discrimination is the answer to the puzzle then if we could find two high school students who are identical except that one is black and one white, we would expect the black high school student to be less likely to apply to and attend college as college is simply less valuable to the black student. To the extent that it is possible to find such hypothetically identical pairs of students, however, the opposite is the case.

Controlling for a wide array of demographics—including *parents' preferences* over their children's educational attainment—black high school students in the U.S. are no less likely, and in some specifications, significantly more likely, to apply to college, to attempt college, and to earn university degrees than their white counterparts.[7] At the same time, (socioeconomic) status *is* a significant positive predictor of educational attainment suggesting that it is inequality, not race or ethnicity, at the root of the problem. [8] This pattern is not consis-

[7]Controlling for parents preferences is meant to account for another common explanation for educational under-attainment: that parents on the bottom rungs of society value education less, and pass this under-valuation on to their children.

[8]A discrete choice model for college application was estimated using the 2002-2006 wave of the Educational Longitudinal Survey (ELS) and the 1988-1994 waves of the ELS, which are large, nationally representative, longitudinal studies that follow American high school students for up to twelve years. The 2002-2006 ELS follows high school students students who were in 10th grade in 2002; the 1988-1994 ELS follows students who were in 8th grade in 1988. The regression specifications control for a wide variety of demographics and

tent with racial (statistical) discrimination, but is consistent with low status undermining confidence. The estimates can be found in the appendix.[9]

Also along these lines, gender is commonly thought to be a status-carrying trait. A link between status and confidence is therefore consistent with findings that women are more reluctant than men to have their earnings based on relative performance criteria (Niederle and Vesterlund (2007)); and that women are less likely than men to ask for raises or promotions or to negotiate their starting salaries or (Babcock and Laschever, 2003). The consequences here are also substantial: Babcock and Laschever estimate that failing to negotiating one's *first* starting salary costs women a loss in lifetime income of $\$500,000$.

Because both the cause and effect here—inequality and confidence, respectively—are especially difficult to identify and disentangle in real-world data, the current paper reports the results of a controlled laboratory experiment which provide direct evidence of a link between status/inequality and confidence in relative performance. In this experiment, inequality was introduced transparently and subjects' beliefs about relative performance were elicited in an incentive-compatible manner. In a nutshell, consistent with the backwards inference hypothesis being on the positive side of inequality boosted confidence, while being on the wrong side of inequality undermined confidence relative to control sessions in which no inequality was introduced. Before discussing the experimental design and the results in more detail, I will briefly discuss the existing evidence of a status-confidence link, and point

attitudes including income, ethnicity, parents' educational background, how much education parents want their children to get, and ability measured by composite test scores.

Strikingly similar patterns are found in both surveys, suggesting a robustness of the observed phenomenon over age and time: none of the minority groups considered were less likely than white, non-hispanic, students to apply to a four-year university up to two years beyond being high-school aged. In fact, some minority groups were significantly more likely to apply to college. On the other hand, applications did vary with socioeconomic status, controlling for various other demographics. In these regressions, socioeconomic status is a composite of parents education, income and occupational prestige, so, controlling for income and parents' education effectively reduces status to parents' occupational prestige. The relationship between socioeconomic status and college achievement has been noted in various other places. In particular, Bowen and Bok, 1998, find a similar pattern using a different index of socioeconomic status providing a robustness check of the regression specification. Also, similar patterns with respect to status are found in the 2002-2006 wave of the ELS when using students' *expectations* of graduating from a four-year university as the dependent variable.

[9] The estimates are meant to be suggestive, rather than exhaustive. All the obvious caveats about omitted variables and misspecified functional forms apply. Obviously, social status in these results can simply be proxying for some other unobserved heterogeneity.

out what the current experiment adds to this body of evidence.

2 Closely Related Literature

One main source of evidence for the general tendency to infer backwards from resource levels, rewards, or other status characteristics (gender, education level, occupational prestige, etc.) is a long history of experiments demonstrating that individuals who rank low on a particular status-carrying characteristic tend to defer to the (conflicting) judgments of individuals who have higher level of the same status characteristic, even when the characteristic is unrelated to the judgment at hand (see, *inter alia*, Moore, Jr. 1968; Stewart and Moore, Jr., 1992; Kalkhoff and Barnum, 2000)). Although this body of experimental results is the source of several distinct veins of research in social psychology—expectation states theory, status characteristics theory and others—the phenomenon itself is typically called status generalization.

In the typical status generalization experiment, one subject, **A**, is paired with a partner, **B**, who is a confederate of the experimenter. **A** is led to believe that **B** has either higher or lower status than B on some dimension—gender, education, etc. The pair is then jointly presented with an estimation task whose correct answer is purposely highly ambiguous.[10] **A** can confer with **B** before submitting her answer; for his part, **B** is instructed to (almost) always disagree. The typical result is that when **A** is led to believe that **B** has higher status than herself, **A** is more likely to defer to the conflicting recommendation of **B** than when **B** has lower status than **A**. Much of the literature takes these deference patterns as *prima facie* evidence of a tendency to infer backwards from status to ability. However, a handful of papers provide direct evidence of this pattern in beliefs (Strodtbeck, et. al, 1957; Harrod, 1980; Wood and Karten, 1986; Oldmeadow, et al., 2003).

A closely-related vein of research examines stereotype-related patterns in in performance, rather than confidence. Beginning with Steele and Aronson, 1995, the main finding

[10] An example of one commonly-used task is to have the pair view two large rectangles composed of smaller black and white rectangles and guess which of the large rectangles contains a greater proportion of white. Of course the slides are constructed to make the proportions virtually indistinguishable.

in this literature is that increasing the salience of an identity widely associated with a stereotype leads to performance patterns consistent with that stereotype. For example, Steele and Aronson found that black students performed worse than which students on what they were led to believe was a test diagnostic of ability when race was made salient. More recently, Hoff and Pandey (2006) use a different set of social identities (caste membership in India) to provide evidence in an incentive-compatible experiment that making a low-status affiliation (caste) salient, in a setting with status differences, generates under*performance* in the low-status group. One explanation for decreased performance is that inequality or discrimination reduces prior expectations of success on the involved tasks which reduces effort. Since these experiments typically involve tasks requiring a good deal of (mental) effort—"hard" math questions, aptitude tests, repeated maze-solving, etc.—this could explain performance differences. So far, the evidence for this explanation of stereotype threat has been mixed mixed (see, the discussion and results in Cadinu *et al*, 2003).

3 Contributions of the Current Experiment

The current experiment contributes to the existing literature in a variety of was. First of all, economists' standard concerns about the lack of incentive compatibility and the use of deception that is common in sociology and social psychology apply here. With few exceptions, the studies just outlined involve deception and lack pecuniary incentives. The current experiment involves pecuniary incentives, and does not involve deception of any kind.

Secondly, all of the studies just mentioned use real-world indicators of status to induce (often fictitious) status differences, making the identification of causality problematic. As the statistical discrimination literature, starting with Arrow, 1972, makes clear, expected correlations between status and ability can be self-fulfilling and make real-world indicators of status actually informative of ability. This is particularly true of the stereotype threat, and related, literature as the status-carrying traits used are those associated with long histories of discriminatory treatment. On the other hand, the current experiment introduces inequality in a simple, transparent and essentially informationless manner.

Thirdly, most of the evidence involves inferring status-related beliefs from deference patterns. However, the observed deference patterns are also consistent with *preferences* for deference, preferences for conformity or aversion to conflict.[11] In those few studies in which beliefs are elicited directly, they are typically *ex-ante, absolute, own*-performance beliefs, making the inference of relative performance beliefs problematic.[12] The current experiment elicits beliefs directly. In particular, I elicit both individual *relative* performance beliefs and group-level relative performance beliefs *ex-post*, allowing for an analysis of the effect of own-confidence on group-level discrimination. Eliciting beliefs *ex-post* rather than ex-ante contributes to the existing literature by providing evidence concerning whether inequality-based confidence patterns survive repeated learning opportunities.[13]

Fourthly, the standard tasks used in the literature are tasks in which subjects have effectively no private information or insight, and therefore have nothing about which to be confident; in which subjects have no way to assess their true relative performance either before or after the experiment; and in which effort plays a large role in performance. On the other hand, in many situations where confidence could have economically-significant consequences individuals have some information about their relative performance/ability, even if this information is very incomplete. For instance, students typically know the distribution of grades in their courses, which may be partially informative of their rank in the pool of applicants to a particular college; Or, workers may observe their own performance, but only observe the performance of the colleagues whose work-schedules overlap with their own.

To address these last concerns, to mimick the real world in the current experiment the task was chosen to provide individuals with feedback concerning their own performance,

[11] The classic conformity study is due to Asch, 1956

[12] See the references above. This also applies, to some extent, to Hoff and Pandey, 2006). Notable exceptions include the early experiments dealing with juror competence, such as Strodtbeck, et. al, 1957.

[13] Of course, both *ex-ante* and interim beliefs could also have been elicited. I chose not to do this in order to lessen the salience of explicit competition, and because the phenomenon of primary interest in this paper hinges upon ex-post performance beliefs.

as well as partial feedback on others' performance. The task chosen was simple enough to implement in a laboratory setting, while complex enough to require analytic ability; and to allow independent measures of confidence in relative performance and an alternate, widely adopted definition of overconfidence: overconfidence in signal precision.

Finally, in the current literature experimentes typically use tasks in which performance depends largely on effort. On the other hand, in the current experiment the task chosen was one in which performance depends more on underlying ability than effort, as eliciting beliefs about relative ability is the ideal goal.[14] Specifically, as detailed below, experimental subjects engaged in multiple rounds of a standard social learning game.

The results clearly demonstrate that inequality, in and of itself, can effect confidence. While actual performance on the task did not vary with inequality, subjects who were on the wrong side of inequality were more likely to perceive their performance to be worse than the majority of other subjects; At the same time, subjects on the positive side of inequality were more likely to believe that they outperformed the majority of other participants. There is also evidence of discriminatory beliefs against low status subjects. The mechanism underlying discrimination was a perceived similarity with in-group members coupled with individual's status-based confidence patterns: subjects perceived their own performance to be representative of their group's performance, resulting in a majority of both high status and low status subjects believing that the low status group performed worse. Finally, there was no evidence of an inequality-induced *preference* for deference, or overconfidence in signal quality.

The remainder of the paper is organized as follows. First, the experimental design is presented in detail. In section 3, results are presented. Section 4 provides a summary of results and concluding remarks.

[14] Although performance in almost any conceivable task entails a bit of both ability and effort.

4 Experimental Design

There were three versions of the experiment. Each subject participated in exactly one of the three versions. In each version, subjects played ten rounds of an urn-guessing game that is common in the social learning literature:

- Nature chooses one of two possible urns with commonly known probability $\frac{1}{2}$.
  - Urn A contains two red balls and one white ball.
  - Urn B contains two white balls and one red ball.
- Each player is randomly matched with (up to) seven other players.
- Within each matching, each player is randomly assigned a sequence number, $t \in \{1 \ldots 8\}$
- According to this sequence number, within each matching each player submits his or her guess about which urn was chosen by nature.
- Specifically, within a particular matching, the t-th player:
  1. observes one private draw, with replacement, from the chosen urn;
  2. observes the $(t-1)$ previous guesses of other players within the matching;
  3. submits one of two guesses: urn A or urn B.

This game is appropriate in the current context since it produces a clear performance measure (percent or number of correct guesses), and performance requires analytic ability but does not depend much on effort. Furthermore, this game allows for natural within-game feedback about one's own and others' performance without making this comparison overly salient. The game also allows the separation of preferences for deference from under- or over-confidence in performance.

The difference between the three versions of the experiment was the payoff structure, summarized in Table 5.2, which is designed to induce salient inequality directly. In the first version of the experiment—Pay Inequality Version(PIV)—half of the subjects in each session were randomly assigned to a "high-paying" group in which correct urn guesses paid \$4 and incorrect guesses paid \$2. The remaining subjects were assigned to a "low-paying" group in which correct guesses paid \$2, while incorrect guesses paid nothing. Subjects' pay groups persisted over all ten rounds of game play. When making their decisions, subjects were informed of the history of previous guesses along with the pay-group affiliation of each preceding guesser.

		Payment per Round (Dollars)	
Version		Correct Guess	Incorrect Guess
PIV	high-pay group	4	2
	low-pay group	2	0
C-HP		4	2
C-LP		2	0

Table 1: Payments, by version

The payoff difference was intended to induce a feeling of either high or low relative status among highly paid and poorly paid subjects, respectively. Therefore, in the analysis that follows, subjects in the PIV version of the experiment will be referred to as either high status or low status subjects. Note that even though there was an average payoff difference associated with differing status levels, the marginal reward for a correct guess was the same in each payment group. The high-paying group effectively received a \$20 lump-sum increase in their participation fee (\$2 $\times$ 10 rounds). Apart from this lump sum compensation difference, the payoff structure was replicates the setup in Anderson and Holt (1997).

The second and third versions of the experiment collectively serve as a control version by implementing the payoff structure for each of the status groups in the PIV version separately. Within each of these control versions, all subjects in each session faced identical payoff structures. In the Control High-Pay version (C-HP), earnings mirrored those for the high status subjects in the PIV version—$4 for correct guess and $2 for incorrect guesses. Earnings in the third version of the experiment—Control Low-Pay (C-LP)—mirrored earnings for low status subjects in the PIV version: C-LP subjects earned $2 for each correct guess, and nothing for incorrect guesses.

In all three versions of the experiment, after ten round of urn-guessing was completed each subject was asked to estimate their relative performance. Specifically, each subject was asked whether they thought they were in the top or bottom half of subjects in terms of the *accuracy* of their predictions. Additionally, in the PIV version of the experiment subjects were asked whether they thought the high-paying group or the low-paying group made more accurate predictions, overall. Subjects were paid one dollar for a correct answer, and zero otherwise.[15]

5 Results

5.1 The Data

The subjects in the experiments were recruited from undergraduates and staff at the University of California, Berkeley. All sessions were conducted in the X-lab facilities at the University of California, Berkeley.[16]

All together, eight sessions were conducted and 152 subjects participated. The PIV version of the experiment was conducted in four of these sessions. In two sessions, the C-LP version of the experiment was conducted and in the remaining sessions the C-HP version of the experiment was implemented. In total, 84 subjects participated in the PIV version of the experiment, 36 subjects participated in the C-LP version of the experiment and 33 subjects participated in the C-HP version.

Multiple rounds of game play produced 1520 observations, 830 of which stemmed from

[15]Notice that, since the questions were about *relative* performance, and subjects knew their own performance—and hence earnings—there is no hedging motive for lying. Also, since the stakes are relatively small, and the choice binary, risk aversion should generate an incentive to strategically misreport beliefs.

[16]The experiment was programmed and conducted with the software z-Tree (Fischbacher 2007)

Percent of Guesses Correct, by Version			
PIV		Control	
High Pay	Low Pay	High Pay	Low Pay
0.65	0.67	0.66	0.68
(0.025)	(0.025)	(0.031)	(0.019)

1. Standard errors in parentheses.

the PIV version of the experiment. All told, the data contain 220 decision sequences, 150 of which are of length 8. The remaining sequences vary in length from 2 to 7, depending on the number of subjects involved in particular sessions. That is, shorter sequences were necessary whenever the number of subjects in a particular session was not divisible by eight.

5.2 Result 1: No Differences in Actual Performance

Using the most transparent measure, performance did not vary with status. Both highly paid and low-paid subject guessed the correct urn about two-thirds of the time. This was true in both the PIV version of th experiment as well as in the control versions (Table 5.2). Furthemore, the null hypothesis that the distribution of the number of correct guesses is equal across status levels cannot be rejected using various standard tests (rank-sum tests yielded $p = 0.710$ in PIV, $p = 0.936$ in control; Chi-squared tests yielded $p = 0.400$ in PIV, $p = 0.170$ in control.). Nor can the null hypothesis that the distribution of the number of correct guesses was the same in the PIV and the control versions be rejected (rank-sum test: $p = 0.595$; $p = 0.564$).

Another way to measure differences in decision quality is to estimate a parametric model. Toward this end, I estimated a logit quantal response equilibrium model with separate decision-quality parameters for high and low status subjects. For details of the basic model as well as a justification of the appropriateness of this model in the context of social learning games, see Goeree *et al.*, 2007. This estimation revealed that, if anything, low status subjects were less error-prone than their high status counterparts. While status had no effect on average performance *within* the treatment or *within* the control versions of the experiment, comparing overall decision quality in the PIV version to decision quality in the control versions reveals weak evidence that introducing a status difference lowered overall decision quality. The model and parameter estimates appear in the appendix.

5.3 Result 2: Status Affects Confidence in Own Performance

Returning to confidence defined as a sanguine assessment of one's relative abilities or performance, I tested for status-based variation in confidence directly by examining the relationship between subjects' estimates of their relative performance, controlling for their actual relative performance. Here, I found significant evidence of the hypothesized status-confidence link.

Specifically, after the urn-guessing phase of the experiment was over but before subjects were informed of how well they, and others, performed, each subject was asked the following question: "compared to others, how accurate were your earn predictions?" Their answers could be either "top 50 percent" or "bottom 50 percent." These beliefs were elicited in an incentive-compatible manner by paying subjects a small amount if their answer turned out to match reality.[17] High status subjects were significantly more likely to have *perceived* themselves as being in the top half of urn-guessers: 85 percent of high status subjects ranked themselves in the top half, while only 60 percent of low status subjects put themselves in this top category ($\chi^2(1) = 6.9165, p = 0.009$). At the same time, in the control versions of the experiment the highly-paid subjects (C-HP) were *less* likely to have perceived themselves as having performed above the median, although the difference was not significant, suggesting that differences in performance perceptions cannot plausibly be attributed to differences in pay levels alone.

Comparing confidence across versions of the experiment implies that high status boosted confidence while low status undermined confidence. One way to see this is to focus on

[17] It must be noted that the median number of correct guesses varies depending on how widely one interprets the reference group "others." The narrowest definition I considered was defined with respect to subjects within each session. I also considered two slightly broader definitions that turned out to be less problematic as they coincided at a value of 60% of guesses correct. These broader medians incorporated either all guesses within each version, or all guesses across all versions. Within the control versions of the experiment (C-LP and C-HP) these distinctions were unnecessary as the various definitions of the median coincided. None of the results below are changed qualitatively by using different definitions of the median.

	Proportion who perceived themselves to be in top 50%:			
	PIV		HP	LP
	High Status	Low Status		
When True Performance Was:				
Above Median	0.95	0.86	0.86	0.95
	(0.045)	(0.078)	(0.078)	(0.045)
Below Median	0.74**	0.33	0.58	0.50
	(0.104)	(0.105)	(0.149)	(0.139)
Overall	0.85**	0.60	0.76	0.78
	(0.056)	(0.077)	(0.076)	(0.070)

1. Standard errors in parentheses.
2. Significance using χ^2 within-row and version.

Table 2: Perceived versus actual performance, by version

subjects who performed below the median. Call this group of subjects "under-performers." In the PIV version of the experiment, 74 percent of high status under-performers perceived themselves to be in the top half, while only one-third of low status under-performers held such a mistaken perception. This difference was significant ($\chi^2(1) = 6.5128, p = 0.011$).[18] At the same time, among under-performers in the control versions of the experiment (C-LP, C-HP) there was no difference in perceived performance across pay levels—about half of the under-performers in each control version thought they had actually performed above the median ($\chi^2(1) = 0.0678, p = 0.795$). The patterns in perceived and actual performance within and across versions are summarized in Table 2.[19]

More formally, I estimated a discrete choice (logit) model using as the dependent variable a dummy taking the value of one if a subject perceived herself as being in the top half of urn-guessers, and zero otherwise. These estimations included controls for subjects' actual performance and the proportion of others' guesses the subject had seen that were correct. Standard errors were clustered at the session level. The estimates confirmed the link between

[18]The difference using the narrower, within-session, definition of the median mentioned in the previous footnote was 75 percent vs. 14 percent for high status vs. low status subjects, respectively ($\chi^2(1) = 7.3044, p = 0.007$).

[19]A seminar participant once suggested that the patterns could be due to the hot-hand fallacy. While the effects may be related, it is not clear what the hot-hand fallacy would predict in this case, which concerns judgmenta about relative performance, whereas the hot-and fallacy typically applies to biases in individuals' judgments about their own performance.

status and confidence. The coeffcent on the dummy variable for the PIV version of the experiment was signficant and negative, indicating that low status lowered confidence in the PIV version relative to the control versions. At the same time, the coefficient on the interaction between being high status and being in the PIV version was positive and significant, indicating that high status increased confidence in the treatment version. The non-significant coefficent on the high-pay dummy alone suggests that the confidence patterns cannot be accounted for by pay differences alone.

5.4 Result 3: Own-Group Discrimination

To the extent that subjects feel a sense of shared fate with their fellow group-members, patterns in own-confidence can extend to patterns in perceived group performance causing a form of own-group discrimination in beliefs.[20] Specifically, if individuals see their own performances and abilities as indicative of their own-groups' performances and abilities, then inferring backwards from their own performances can lead to biased judgments about their group as a whole. To test for this additional group-level implication of the status-confidence link, after all ten rounds of the social learning game were complete subjects in the PIV version of the experiment were asked to predict which group performed better. Specifically, after subjects were asked to predict whether they, personally, were among the most or least accurate urn-guesses, they were also asked: "Which group got the highest percentage of urn predictions correct?" Subjects earned an additional dollar for a correct answer, and zero additional dollars otherwise.[21]

As expected, subjects' beliefs about their own performance was a significant predictor of their beliefs about their groups' performance, providing a clear indication of a sense of shared fate. Overall, exactly two-thirds of PIV subjects who believed they performed in the top half, personally, also believed their own group performed better than the other group.

[20]This sense of shared fate can occur if the groupings in the experiment come to constitute a social idenity, for instance. There is considerable evidence of this effect. For a discussion, see, e.g., Akerlof and Kranton, 2005 or Haslam, 2004.

[21]Subjects were not allowed to guess that each group performed exactly equally well. Since this would be a zero-probability event, and belief elicitation was incentive compatible, it is my view that omitting this possible answer should not be problematic.

Dependent Variable: Subject Perceived Self in Top Half.				
	(1)	(2)	(3)	(4)
High Pay	-0.209	-0.0367	-0.0863	0.323
	(0.476)	(0.480)	(0.475)	(0.478)
PIV-version	-0.817***	-0.982***	-0.971***	-1.001***
	(0.158)	(0.274)	(0.276)	(0.324)
PIV×(High Pay)	1.800**	1.906**	1.947**	1.676**
	(0.796)	(0.825)	(0.808)	(0.812)
Actually in Top Half	2.244***	0.179		
	(0.521)	(1.348)		
Num Correct Guesses		0.122		
		(2.576)		
(Num Correct Guesses)^2		-0.0799		
		(0.469)		
(Num Correct Guesses)^3		0.0151		
		(0.0259)		
Num Correct Guesses=3			-3.110***	-4.028***
			(1.148)	(1.147)
Num Correct Guesses=4			-3.603***	-3.954***
			(1.169)	(1.316)
Num Correct Guesses=5			-3.194***	-3.451***
			(1.008)	(1.129)
Num Correct Guesses=6			-2.412***	-2.557***
			(0.791)	(0.918)
Num Correct Guesses=7			-1.231	-1.322
			(1.115)	(1.214)
Prop. Others Seen Correct				-141.0*
				(72.00)
(P.O.S.C.)^2				229.1*
				(120.4)
(P.O.S.C.)^3				-121.2*
				(63.53)
Constant	0.212	-0.559	3.016***	31.26**
	(0.150)	(4.102)	(0.803)	(13.20)
Observations	152	152	136	136
Pseudo-R^2	0.211	0.256	0.207	0.236

1. Each estimate involves a logit model of discrete choice.
2. Robust standard errors in parentheses, clustered at session level.
*** p<0.01, ** p<0.05, * p<0.1
3. Models 3 and 4 are restricted to subjects who had an intermediate number of correct guesses, i.e., between 3 and 8 out of 10, inclusive. This is because there was no variation in beliefs among subjects with either very high or very low numbers of correct guesses.

Among PIV subjects who believed they had performed below the median, the figure was about one-third. A chi-square test rejected the null hypothesis that perceived own-rank and perceived group-rank were independent ($\chi^2(1) = 6.9310, p = 0.008$). These patterns are summarized in Table 3. At the same time, subjects' perceived connection between their own-performance and their groups' performance was illusory. A chi-square test failed to reject the null hypothesis that subjects' perceived own-performance was independent of whether their group was actually the better-performing group ($\chi^2(1) = 0.0314, p = 0.859$).

Also in line with expectations, the connection subjects perceived with their group led to a form of discrimination in beliefs: a majority of both pay-groups reported believing that the highly-paid group performed better. Thus, on average subjects perceived a positive correlation between pay and ability, or relative performance. Coupled with the ubiquitous judgment that performance and pay *should be* positively related, backward inferences about performance serve to justify initial group-level inequality.[22]

An interesting question is whether discrimination in beliefs would lead to discrimination in pay. Although this is beyond the scope of the current paper, a 1975 paper by sociologist Cook presents experimental evidence that, when given the opportunity, subjects implement unequal pay schemes for equal work on the basis of inferred ability. Subjects in Cook's experiment who were led to believe they had higher ability than an anonymous partner, and were given the chance to implement either an unequal pay scheme (either in their favor, or their partner's favor) or an equal pay scheme, implemented pay schemes more biased toward themselves than did subjects who were led to believe they had lower ability than their partner.

Estimating a logit discrete choice model allows for the adjustment of estimated standard errors to account for the fact that observations may not be independent within sessions. The estimations, presented in Table 4, yielded the same overall patterns. However, accounting for possible session-level dependence suggests that the correlation between perceived own-

[22] As evidence of the ubiquity of support for this position, notice that, for example, approximately 40% of respondents to the International Social Survey (Question: V52) report that relative performance *should be* the most important determinant of pay difference between people performing similar tasks.

Perception of Own Performance	Own Group Perceived Better		
	High Status	Low Status	Overall
Top Half	0.71*	0.60	0.67***
	(0.077)	(0.100)	(0.061)
Bottom Half	0.33	0.35	0.35
	(0.211)	(0.119)	(0.102)
Overall	0.65	0.50	
	(0.075)	(0.078)	

1. Values reported are the proportion of subjects in each category who reported that their own group performed better.
2. Standard errors in parentheses
3. Significance is with respect to within-column “Bottom Half” value.

Table 3: Perceived Own and Group Performance (PIV version)

performance and perceived group performance did not vary significantly with pay-level. Apparently, individuals felt connected with their group’s outcome regardless of inequality.

This bond has the potential to amplify the effects of inequality within a society. Majority group members who perceive themselves to be personally successful come to believe that their own success is evidence of their group’s superiority. This belief may also come to be accepted by the members of minority groups who perceive themselves to be unsuccessful. Because of confidence patterns, the combination of these two groups can be quite large —40% of low-paid subjects in the current experiment perceived themselves to be under-performers.

5.5 Result 4: No Differences in Signal-Confidence

An alternate definition of (over)confidence that has gained wide acceptance relates to the the perceived precision of one’s private signal (see, e.g., Nöth and Weber, 2003). Since subjects know their signal precision, this type of confidence is unlikely to play a role here. Accordingly, I found little evidence that status-based variation in over- or under-confidence.

Specifically, there was little evidence that pay differences affected subjects’ reluctance to ignore their own signals. To show this, I restricted attention to early-movers where the relatively small set of possible information histories permits meaningful comparisons across

Dependent Variable: Subject Believed Own Group Performed Better.				
	(1)	(2)	(3)	(4)
Top-Half-Belief	1.322***	1.012***	1.044***	1.035***
	(0.150)	(0.376)	(0.385)	(0.368)
HP		-0.0870	-0.386	-0.321
		(0.744)	(0.928)	(1.010)
HP×(Top-Half-Belief)		0.598	0.709	0.756
		(1.275)	(1.351)	(1.287)
Actual-Own-Group-Better			-0.430*	-0.336
			(0.242)	(0.412)
A.O.G.B.×HP				-0.203
				(0.429)
Constant	-0.629**	-0.606***	-0.308*	-0.372*
	(0.296)	(0.170)	(0.160)	(0.194)
Observations	83	83	83	83
R^2	0.061	0.069	0.075	0.075

1. Robust standard errors in parentheses, clustered by session.
2. Top-Half-Belief is a dummy indicating whether a subject believed he or she was in the top half, personally.
3. Actual-Own-Group-Better is a dummy indicating whether a subject's group was actually more accurate.

Table 4: Perceived Own and Own-Group Performance, Logit Model (PIV version)

pay levels. Among this subset of the data, there is no significant pay-level-related variation in subjects' propensity to ignore their own signals. This is true both overall, and when further restricting attention to only those instances where decision-makers private signal are contrary to the history of predecessors' guesses. Average propensities to ignore own signals are presented in Table 5.

		PIV		Control	
		High-Pay	Low-Pay	High-Pay	Low-Pay
Sequence Order					
1		0.13	0.15	0.06	0.02
		(0.050)	(0.042)	(0.034)	(0.020)
2	Overall	0.16	0.12	0.08	0.14
		(0.047)	(0.043)	(0.039)	(0.050)
	Contrary History[1]	0.24	0.20	0.13	0.22
		(0.087)	(0.074)	(0.085)	(0.082)
3	Overall	0.18	0.17	0.24	0.18
		(0.051)	(0.052)	(0.061)	(0.055)
	Contrary History[2]	0.54	0.50	0.59	0.69
		(0.144)	(0.129)	(0.123)	(0.133)

1. Contrary history means the data were restricted to observations in which the predecessor's urn guess contradicts the decision-maker's private signal.
2. Contrary history for sequence order 3 restricts attention to observations in which both predecessors guessed the same urn, and this common guess was contrary to the decision-maker's private signal.
3. Standard errors in parentheses.

Table 5: Propensity to Ignore Own Signal

5.6 Result 5: No Evidence for a Preference for Deference

By further subdividing the the data by the pay level of predecessors, one can shed light on whether there is a preference for deference involved. If so, we would expect low-paid subjects to be more willing to follow the judgments of highly-paid predecessors than low-paid predecessors. Unfortunately, this extra refinement permits only the examination of sequence order two. Given this caveat, the data suggest that there is no pronounced pattern of deferring to a high-status predecessor and, if anything, low status subjects are *more* reluctant to ignore their signals following the guesses of high status predecessors than low

Propensity to Ignore Own Signal.			
		Predecessors' Group	
	DM's Group		
sequence order			
		HP	LP
2	HP	0.11	0.31
		(0.111)	(0.120)
	LP	0.11	0.24
		(0.111)	(0.095)

1. Standard errors in parentheses.
2. Restricted to observations where own-signal contradicts history of guesses.
3. Restricted to PIV version.

Table 6: Propensity to Ignore Own Signal, by Predecessor's Group

status predecessors. The patterns are presented in Table 6.

6 Summary and Conclusion

The results provide direct evidence of a link between inequality and confidence. The patterns in the data suggest that being on the wrong side of inequality can undermine confidence and that being on the favorable side of inequality can boost confidence.[23] Through its effect on confidence, therefore, inequality can be self-sustaining even without long patterns of external discrimination or mistreatment.

A second pattern in the data relates to the spontaneous creation of discrimatory beliefs. Inequality-induced patterns in individuals' confidence patterns combined with a perceived in-group similarity led a majority of both high status and low status subjects to believe that low status subjects performed worse. This is despite the fact that the correlation between own and own-group performance was illusory.[24]

[23]While neither group of subjects was made particularly *under*confident by inequality, this may have had much to do with the pronounced overconfidence of the subject pool as can be inferred from the control sessions. In light of the current experiment, this general overconfidence may be due to the subject pool involved—students at a highly prestigious research university who essentially won a tournament to achieve this position.

[24]A glimmer of hope can be seen in the fact that low status subjects who perceived themselves to be above-median performers also believed that their group performed better; this is consistent with wanting highly successful members of one's group to represent the group, which can be seen in, e.g., the preponderence of wealthy members of the U.S. congress.)

Together, these patterns in confidence provide an explanation for self-reinforcing inequality that does not rely on expectations of external discrimination, or past histories of severe discrimination. Simple, transparent salient inequality leads low status individuals to effectively discriminate against themselves. In particular, even in a perfect world in which all external discrimination is eliminated, inequality can still be self-reinforcing. The poor, as the saying goes, will always be with us.

A Tables and Figures

Dependent Variable: Applied College Within Two Years of Typical High School Graduation Date.

	2002 ELS				1988 ELS			
	Odds Ratio	t	Prob(t)	Ln(Odds Ratio)	Odds Ratio	t	Prob(t)	Ln(Odds Ratio)
Socioeconomic Status (B.Y.)								
2nd Quartile	1.101	1.185	0.237	0.096	1.172	1.89	0.066	0.159
3rd Quartile	1.431	3.425	0.001	0.359	1.600	5.151	0.000	0.47
Highest Quartile	1.956	5.129	0.000	0.671	2.117	5.235	0.000	0.75
vs. Lowest Quartile								
Male	0.52	-11.088	0.000	-0.655	0.756	-4.607	0.000	-0.28
Student's Race/Ethnicity								
Asian	1.631	4.002	0.000	0.489	1.896	3.825	0.000	0.639
Black	1.346	2.983	0.003	0.297	1.970	6.49	0.000	0.678
Hispanic	1.084	0.842	0.401	0.081	1.202	1.424	0.162	0.184
Multi	0.962	-0.266	0.791	-0.039				
vs. White, non-Hispanic								
Family Income (B.Y.)								
10K-35K	1.157	1.169	0.244	0.146	1.275	1.519	0.137	0.243
35K-75K	1.179	1.237	0.218	0.164	1.358	1.722	0.093	0.306
75K-200K	1.342	1.942	0.054	0.294	2.328	3.45	0.001	0.845
$>$ 200K	2.587	3.613	0.000	0.951	3.829	4.201	0.000	1.343
vs. below 10K								
Standardized Test (B.Y.)								
2nd Quartile	1.569	5.166	0.000	0.451	1.843	4.687	0.000	0.611
3rd Quartile	2.471	10.909	0.000	0.905	2.853	9.251	0.000	1.048
Highest Quartile	4.492	15.639	0.000	1.502	6.354	15.091	0.000	1.849
vs. Lowest Quartile								
Parents' Education (B.Y.)								
B.A. or More	0.991	-0.096	0.924	-0.009	1.543	4.799	0.000	0.434
vs. $<$ B.A.								
Parents' Wishes (B.Y.)								
BA or More	1.426	4.187	0.000	0.355	2.117	8.852	0.000	0.75
vs. $<$ B.A.								
Student's Expectations (B.Y.)								
B.A. or More	2.45	10.744	0.000	0.896	2.633	10.441	0.000	0.968
vs. $<$ B.A.								
Intercept	0.426	-6.786	0.000	-0.854	0.044	-20.029	0.000	-3.13
Pseudo R^2	0.147				0.273			
Percentage of Cases :	77.29%				75.30%			

1. Coefficients are from a logistic regression, using a dummy dependent variable taking the value of 1 if the respondent had applied to at least one 4-year university within two years of the age at which he or she (should have) graduated high school.
2. Source: 2002-2006 and 1988-1994 wave of the U.S. Educational Longitudinal Survey (ELS).
3. Estimates constructed using the National Center for Education Statistics on-line data analysis system (DAS).

Proportion perceiving themselves in the top 50%, by number correct						
	Treatment			Control		
# correct	HP	LP	(# HP, # LP)	HP	LP	(# HP, # LP)
2	—	0.00	(0,1)	—	—	(0,0)
3	0.5	1.00	(2,1)	0.33	—	(3,0)
4	0.5	0.00	(2,1)	0.50	—	(2,0)
5	0.71	0.00	(7,3)	0.33	0.60	(3,5)
6	0.88	0.41	(8,15)	1.00	0.44	(4,9)
7	0.89	0.86	(9,7)	0.70	0.92	(10,12)
8	1.00	0.80	(9,10)	1.00	1.00	(6,8)
9	1.00	1.00	(4,2)	1.00	1.00	(5,2)
10	—	1.00	(0,2)	—	—	(0,0)
Overall	0.85	0.60	(41,42)	0.76	0.78	(33,36)
Actual	0.54	0.50		0.64	0.61	

1. Because of the lumpy nature of the performance measure, the "top 50%" comprises more than half the observations. In the treatment sessions it comprises 51.81%, while in the control sessions it involves 62.32%.

B A 2-parameter Quantal Response Equilibrium model

To get a single, summary, measure of decision quality it is helpful to estimate a parametric model of decision-making. One standard model that has been used for this purpose using data from social learning experiments is a logit quantal response equilibrium (QRE) model. For a formal derivation of the model, see Goeree *et al.*, 2007.

The gist of the model is that individuals' guesses are subject to a commonly known error probability. The first person in a decision sequence guesses "urn A" or "urn B" according to the expected utility of each of these guesses, plus a random error term, u_i. The logit QRE model assumes that errors are distributed according to the extreme value, type I, distribution. Since the error distribution is common knowledge, the second-mover can estimate the probability of each urn being chosen given the first person's guess and his or her own private draw. The second-mover uses these probabilities to calculate the expected utility of guessing urn A versus guessing urn B, and chooses according to these expected utilities—again subject to a random error. In the same way, each subsequent mover can form expectations about the correct urn and makes their own decision. In general, this leads to a one-parameter choice function that can be estimated from the data.

$$Prob(Guess = A|draw = a) = \frac{e^{\lambda * u(Guess=A|draw=a)}}{e^{\lambda * u(Guess=A|draw=a)} + e^{\lambda * u(Guess=B|draw=a)}}$$

In this choice function, λ captures the precision, or rationality, of choice. As $\lambda \to 0$, choices become completely random; as $\lambda \to \infty$ the probability of making the Bayesian-correct choice approaches 1.

The 2-parameter model estimated here simply assumes that high-paid and low-paid subjects have different, commonly known, error terms: $\lambda_h and \lambda_l$, respectively. Comparing these estimated parameters gives an indication of whether pay/status differences affected the quality of decision-making. The estimated parameters are given in the table below. By this measure, pay differences had no significant effect on decision quality: in both the control and the PIV versions of the experiment, the estimated decision-quality parameters statistically

Estimated Decision Quality, QRE Model.		
	$\underline{\lambda_h}$	$\underline{\lambda_l}$
PIV Version	2.02	2.15
	(0.160)	(0.224)
Control Versions	2.47	3.05
	(0.222)	(0.400)
1. Standard errors in parentheses.		

indistinguishable. In fact, if pay difference had any effect with version of the experiment, it seems to have been increasing the decision quality of the less-well-paid subjects. Comparing the estimated coefficients across versions, however, reveals weak evidence that introducing salient inequality reduced the overall quality of decision-making for all subjects.

C Induced Status Differences

There was evidence in post-experiment surveys that payoff differences in the PIV version induced a status difference among subjects, while no such status difference existed between the control versions of the experiment. Specifically, after the experiment was completed and subjects were informed how many questions they had guessed correctly, each subject was asked: "If an impartial observer were to rank everyone they knew, from worst to best, using whatever criteria they wanted to, where do you think they would place you?" Subjects responded on a scale of one to seven with one corresponding to "among the worst" and seven corresponding to "among the best."[25]

To estimate the difference in perceived status levels between high status and low status PIV subjects, I regressed subjects' numerical response to the above question on a dummy variable that took the value of one if the subject was assigned to the high status group. Since subjects knew their actual relative performance at the time they answered the survey questions, I included subjects' (normalized) rank in the regression.[26] The resulting estimate

[25] Although the concept of status is hard to define precisely, preeminent social psychologist Roger Brown, in his classic work *Social Psychology* 1965, treats the answer to the hypothetical question "who are the better people in your community and who are the lesser ones" as synonymous with the concept of status.

[26] Subjects' actual rank in session, $r \in \{1, \ldots, \text{number of subjects in session}\}$, was converted into a percentile to make the values directly comparable across sessions.

suggested that high status subjects in fact thought that an outside observer would view them as superior to low status subjects: the average response from high status subjects was significantly higher ($p = 0.029$).[27]

To test whether this effect was solely due to earnings differences, I performed the same estimation pooling subjects in the C-LP and C-HP versions of the experiment. Here, the dummy for high status took the value of one if the subject had participated in the C-HP version and zero if the subject had participated in the C-LP version. The estimated coefficient on the dummy variable was not significantly different from zero ($p = 0.270$), suggesting that earnings differences alone cannot explain the perceived status differences.

References

[1] George A. Akerlof and William T. Dickens. The economic consequences of cognitive dissonance. *The American Economic Review*, 72(3):307–319, 1982.

[2] Gordon Willard Allport. *The Nature of Prejudice*. Addison-Wesley Publishing Co., 1954.

[3] Lisa R. Anderson and Charles A. Holt. Information cascades in the laboratory. *The American Economic Review*, 87(5):847–862, December 1997.

[4] K. Arrow. *Racial Discrimination in Economic Life*, chapter Some Models of Race in the Labor Market. D.C. Heath, Lexington, MA, 1972.

[5] S. Asch. Studies of independence and conformity: A minority of one against a unanimous majority. *Psychological Monographs*, 70(416), 1956.

[6] Linda Babcock and Sara Laschever. *Women don't ask : negotiation and the gender divide*. Princeton University Press, Princeton, NH, 2003.

[27] The coefficient on the high status dummy variable was 0.58. It should be noted that there were four subjects in the PIV version of the experiment whose responses were lost because of a technical malfunction. There is no reason to suspect this malfunction affected one group more than another, however.

[7] Gary S. Becker. *The Economics of Discrimination, 2nd ed.* University of Chicago Press, Chicago, Il, 1971.

[8] William G. Bowen and Derek Bok. *The Shape of the River: Long-term Consequences of Considering Race in College and University Admissions*. Princeton University Press, 1998.

[9] Roger Brown. *Social Psychology.* The Free Press, New York, 1965.

[10] Colin Camerer and Dan Lovallo. Overconfidence and excess entry: An experimental approach. *The American Economic Review*, 89(1):306–318, 1999.

[11] Karen S. Cook. Expectations, evaluations and equity. *American Sociological Review*, 40(3):372–388, 1975.

[12] Urs Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2):171–178, 2007.

[13] Jacob K. Goeree, Thomas R. Palfrey, Brian W. Rogers, and Richard D. McKelvey. Self-correcting information cascades. *Review of Economic Studies*, 74(3):733–762, July 2007.

[14] Wendy Jean Harrod. Expectations from unequal rewards. *Social Psychology Quarterly*, 43(1):126–130, 1980.

[15] Karla Hoff and Priyanka Pandey. Economic consequences of social identity: Discrimination, social identity and durable inequalities. *AEA Papers and Proceedings*, pages 206–211, May 2006.

[16] Moore Jr. James C. Status and influence in small group interactions. *Sociometry*, 31(1):47–63, March 1968.

[17] Oliver Sheldon John T. Jost, Brett W. Pelham and Bilian Ni Sullivan. Social inequality and the reduction of ideological dissonance on behalf of the system: evidence of

enhanced system justification among the disadvantaged. *European Journal of Social Psychology*, 33:13–36, 2003.

[18] John T. Jost, Brett W. Pelham, and Mauricio R. Carvallo. Non-conscious forms of system justification: Implicit and behavioral preferences for higher status groups. *Journal of Experimental Social Psychology*, 38(6):586 – 602, 2002.

[19] Will Kalkhoff and Christopher Barnum. The effects of status-organizing and social identity processes on patterns of social influence. *Social Psychology Quarterly*, 63(2):95–115, June 2000.

[20] Ulrike Malmendier and Geoffrey Tate. Ceo overconfidence and corporate investment. *The Journal of Finance*, 60(6):2661–2700, 2005.

[21] Ulrike Malmendier and Geoffrey Tate. Who makes acquisitions? ceo overconfidence and the market's reaction. *Journal of Financial Economics*, 89(1):20 – 43, 2008.

[22] Sara Frigerio Lisa Impagliazzo Mara Cadinu, Anne Maass and Samira Latinotti. Stereotype threat: the effect of expectancy on performance. *European Journal of Social Psychology*, 33(2):267–285, 2003.

[23] Don A. Moore and Daylian M. Cain. Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes*, 103(2):197 – 213, 2007.

[24] Muriel Niederle and Lise Vesterlund. Do women shy away from competition? do men compete too much? *Quarterly Journal of Economics*, 122(3):1067–1101, August 2007.

[25] Julian A. Oldmeadow, Michael J. Platow, Margaret Foddy, and Donna Anderson. Self-categorization, status, and social influence. *Social Psychology Quarterly*, 66(2):138–152, 2003.

[26] Cecilia L. Ridgeway, Elizabeth Heger Boyle, Kathy J. Kuipers, and Dawn T. Robinson. How do status beliefs develop? the role of resources and interactional experience. *American Sociological Review*, 63(3):331–350, 1998.

[27] Richard Roll. The hubris hypothesis of corporate takeovers. *Journal of Business*, 59(2):197–216, April 1986.

[28] C.M. Steele and J. Aronson. Stereotype threat and the intellectual test performance of the african-americans. *Journal of Personality and Social Psychology*, 69:797–811, 1995.

[29] Penni A. Stewart and James C. Moore Jr. Wage disparities and performance expectations. *Social Psychology Quarterly*, 55(1):78–85, 1992.

[30] Fred L. Strodtbeck, Rita M. James, and Charles Hawkins. Social status in jury deliberations. *American Sociological Review*, 22(6):713–719, 1957.

[31] Wendy Wood and Stephen J. Karten. Sex differences in interaction style as a product of perceived sex differences in competence. *Journal of Personality and Social Psychology*, 50(2):341–347, Feb 1986.