

Running Head: GENDER DIFFERENCES IN SELF-HANDICAPPING

She Works Hard for the Money: Valuing Effort Underlies  
Gender Differences in Behavioral Self-Handicapping

Sean M. McCrea

Universität Konstanz

Edward R. Hirt

Bridgett J. Milner

Indiana University-Bloomington

Final published version:

**Journal of Experimental Social Psychology**  
Volume 44, Issue 2, March 2008, Pages 292-311  
[doi:10.1016/j.jesp.2007.05.006](https://doi.org/10.1016/j.jesp.2007.05.006)



### Abstract

Research in the area of self-handicapping has consistently demonstrated a robust yet puzzling gender difference in the use of and evaluation of behavioral self-handicaps; women (1) are less likely to use these forms of handicaps, particularly those involving the actual or reported reduction of effort, and (2) evaluate the use of these handicaps by others more negatively than do men. The present research examines several possible explanations for these consistent gender differences and finds that the personal value placed on effort is an important mediator of these effects.

Key Words: Self-handicapping, Gender differences, Individual differences, Personality

## She Works Hard for the Money: Valuing Effort Underlies Gender Differences in Behavioral Self-Handicapping

Self-handicapping involves creating or claiming an obstacle prior to a performance in order to provide a viable excuse for possible failure (Jones & Berglas, 1978). In this manner, the individual can protect self-esteem and impressions of ability (McCrea & Hirt, 2001). Research has identified a wide range of behaviors individuals use in this manner, and numerous situational factors and individual differences that moderate self-handicapping. Some variables, such as increased task importance (Sheppard & Arkin, 1989), ego-relevance (Pyszczynski & Greenberg, 1983), and the presence of audiences (Hirt, McCrea, & Kimble, 2000) increase self-handicapping by making the performance more important and threatening to the individual. Others, such as having uncertain self-esteem (Harris & Snyder, 1986), or experiencing noncontingent success (Berglas & Jones, 1978) increase this behavior by leading the individual to feel less confident of his or her ability to perform well. These findings fit well within self-handicapping theory. However, one of the most consistent individual differences in self-handicapping behavior has proven particularly difficult to explain. Research has repeatedly found a gender difference that men appear to self-handicap to a greater degree than do women, yet the reason for this difference remains elusive (Arkin & Oleson, 1998; Rhodewalt, 1990). The primary goal of the present research was to provide an explanation for this intriguing gender difference.

Researchers on self-handicapping have distinguished between claimed self-handicaps, in which the individual merely states that an obstacle to performance exists, and behavioral handicaps, in which the individual actually creates obstacles to performance (Leary & Shepperd, 1986). This distinction has proven critical in terms of the gender difference in self-handicapping. Hirt, Deppe, and Gordon (1991) made both types of handicaps available to participants prior to

an exam in order to examine which would be preferred. Although both men and women engaged in claimed self-handicapping (i.e., reported stress), only men behaviorally handicapped (i.e., withdrew practice effort). Other research has supported this finding using other types of behavioral handicaps, including drug use (Berglas & Jones, 1978) and choice of performance setting (Rhodewalt & Davison, 1986). Indeed, many studies involving behavioral self-handicapping have included only male participants (e.g., Kolditz & Arkin, 1982; Rhodewalt, Saltzman, & Wittmer, 1984). Of the studies that have directly compared behavioral self-handicapping among men and women, the vast majority (e.g., Berglas & Jones, 1978; Harris & Snyder, 1986; Hirt et al., 1991; Hirt et al., 2000; Rhodewalt & Davison, 1986), conducted by a range of researchers and using a variety of methodologies, have found that only men behaviorally handicapped (for exceptions see Ferrari & Tice, 2000; Strube & Roemmele, 1985). On the other hand, nearly every study investigating claimed self-handicaps has found that they are used by women as well as men (Arkin & Oleson, 1998; Rhodewalt, 1990). Thus, the overall pattern of results strongly suggests that women do not usually behaviorally self-handicap. Indeed, in a review of these studies, Rhodewalt (1990) stated that this gender difference is one of “the most consistent findings” in the self-handicapping literature.

#### *Explanations for the gender difference*

Several explanations for the gender difference have been put forth in the past, focusing on the notion that men are in some way more motivated to self-handicap. However, to date, there has been a relative paucity of direct attempts to explore these possibilities, and studies conducted so far have only served to rule out possible explanations, rather than provide direct evidence for any particular explanation.

One possibility that has been considered is that women may be less threatened by task failure (Harris, Snyder, Higgins, & Schrag, 1986; Snyder, Ford, & Hunt, 1985) or that women may not be as apt to self-handicap when called upon to publicly display a desired ability (Rhodewalt, 1990). If this were true, women may not choose to engage in behavioral self-handicapping unless the threat to self-concept and/or public impression is increased (Hirt et al., 1991). A recent study (Hirt et al., 2000) tested this explanation by manipulating public self-focus via the presence of a camera and measured practice effort prior to a test of intelligence. Public self-focus did increase evaluative concern and uncertainty more for men than for women. However, whereas controlling for such concern completely explained the reduction in men's practice effort in the public self-focus condition relative to a control condition, concern did not significantly predict women's practice behavior. Furthermore, controlling for concern did not reduce the effects of gender on self-handicapping within the public self-focus condition. Hirt et al. (2000) also reported that gender differences in behavioral self-handicapping were not explained by participants' rating of the importance of the performance. A subsequent replication of this study (Koch, Hirt, & McCrea, 2003) showed that increased public self-focus resulted in increased claimed self-handicapping for both men and women, suggesting it is the type of handicap, not the level of threat to self or the presence of public scrutiny, that underlies this gender difference.

A second potential explanation stems from the fact that men tend to report higher self-esteem than do women (Feingold, 1994; Kling, Hyde, Showers, & Buswell, 1999). Perhaps those high in self-esteem are more prone to engage in behavioral self-handicapping because they have more to lose in performance situations. However, increased self-handicapping has been demonstrated both among those high (McCrea & Flamm, 2007; McCrea & Hirt, 2001; Tice, 1991), and low in self-esteem (Strube & Roemmele, 1985; Tice, 1991). Moreover, controlling for self-esteem does

not appear to affect the gender difference (Harris & Snyder, 1986; McCrea & Flamm, 2007; McCrea & Hirt, 2001).

A third possibility is that women do not behaviorally self-handicap in the stereotypically masculine academic domain (cf., Swim & Sanna, 1996), because the importance of an academic performance is lower for women. Thus, this explanation argues that the locus of the effect resides in the performance domain being examined, and predicts that women would behaviorally self-handicap in a performance domain in which they are expected to excel. However, research examining self-handicapping in the more stereotypically feminine domain of social interaction has still found that men behaviorally handicap more (Hirt, 1993; Kimble, Funk, & DaPolito, 1990). Moreover, it is the case that women make use of claimed handicaps even in the academic domain (Hirt et al., 1991; Koch et al., 2003; Rhodewalt, 1990). Thus, it is unlikely that gender differences are limited to the academic domain, although more research is required to definitively rule out this interpretation.

Given that self-handicapping is motivated in part by impression management concerns (cf., Kolditz & Arkin, 1982), we have also considered the possibility that women are not afforded the same attributional benefits for self-handicapping as are men. Past work has shown that failure by women is more likely to be attributed to lack of ability, whereas failure by men is attributed to lack of effort, at least for masculine tasks (Dweck, Davidson, Nelson, & Enna, 1978; Swim & Sanna, 1996). Therefore, it could be that women do not behaviorally handicap because they expect observers will blame their failures on lack of ability rather than lack of effort.

To test this possibility, Hirt, McCrea, and Boris (2003) conducted a study in which men and women read a short vignette about Chris, a student with an important upcoming test (see also Luginbuhl & Palmer, 1991). In these vignettes, the gender of the target (Chris) was varied. Chris

was described as concerned about the test but was not prepared; his/her preparation the night before the exam was manipulated. Chris either went to a movie, staying out for the entire evening (foregoing the opportunity to study), or (s)he stayed home to study. Thus, in the movie condition, Chris' behavior was designed to be consistent with behavioral self-handicapping.<sup>1</sup>

Contrary to the prediction that women are not afforded the same attributions for self-handicapping as are men, no effect for target gender was found. Rather, participant gender played a crucial role in determining reactions to the target. Women were more critical of the self-handicapping target (male or female) than men on both ability-related and interpersonal dimensions. Subsequent replications revealed that men were willing to give the self-handicapper the "benefit of the doubt" when an alternative explanation for the behavior was available (e.g., target felt peer pressure to go out), whereas women did not. Furthermore, these gender differences were mediated by the perceived motives of the target. Women were more likely than men to view the target's self-handicapping as indicative of a stable character defect, such as dispositional laziness or a lack of self-control. Moreover, only women endorsed the possibility that the target was motivated by an ulterior motive related to self-handicapping (i.e., making excuses in case they failed or trying to showoff in the event of success). Thus, women were suspicious of individuals who did not put forth effort for the exam, and more likely to view the behavior as self-handicapping or reflecting some sort of character flaw. These findings led Hirt et al. (2003) to suggest that gender differences in audience reactions to behavioral self-handicapping result from women placing increased importance on effort compared to men. As a result, women react negatively to someone who does not try their best. Furthermore, women may not engage in these types of self-handicapping behaviors themselves, not because they are less motivated to do so, but because they view certain types of self-handicaps negatively.



*Gender differences in the valuing of effort*

A review of the self-handicapping literature thus led us to hypothesize that differences in the perceived costs of behavioral handicaps (rather than differences in evaluative concern, self-esteem, importance of the performance, domain of the performance, or likely audience reactions to male and female targets) might serve as the root cause of these gender differences. Clearly, handicaps appear to involve a tradeoff of protecting ability beliefs and perceptions of underlying competence in exchange for numerous costs, including performance decrements and impressions of irresponsibility, complaining, or excuse-making (Baumeister & Scher, 1988). We suggest that for women, the perceived costs of behavioral self-handicapping are simply too high because they place more value on putting forth effort than do men. Conversely, men appear willing to use these handicapping strategies because, for them, these costs do not outweigh the advantages of protecting conceptions of ability. Gender differences in audience reactions to reduced effort similarly reflect the differential weight men and women place on putting forth effort. Thus, women may come to disapprove of a lack of effort by others as well as in their own behavior, and therefore do not self-handicap in this manner.

We readily acknowledge that we are not the first to argue that women value effort more than men. Indeed, the claim that women generally put forth more effort and value effort more than men has received support from numerous studies. Female students have more effortful study strategies, report studying harder, and are more likely to adopt effortful learning goals than male students (e.g., Ablard & Lipschultz, 1998; Stricker, Rock, & Burton, 1993; Zimmerman & Martinez-Pons, 1990). Men procrastinate more on important tasks, such as completing one's psychology research requirement, than do women (e.g., Cooper, Baumgardner, & Strathman, 1991). Men also tend to under-perform and women over-perform in college academic

performance when controlling for standardized test score. Moreover, this pattern is at least in part due to differences in reported study effort (Stricker et al., 1993). However, the present research represents the first attempt to explain the robust gender differences in behavioral self-handicapping in terms of differences in the value placed on effort. Specifically, the current set of studies sought to directly test for gender differences in the importance and value placed on effort and to examine whether these beliefs can potentially explain the gender difference in both (1) the use of behavioral self-handicapping as well as (2) reactions to targets who engage in behavioral self-handicapping.

### The Present Studies

To investigate this possibility, we began a line of research with the goal of developing a set of measures to assess the extent to which individuals value effort. Indeed, although there is good consensus that men and women differ in the value they place on effort, there are many different facets of this relationship that could potentially mediate gender differences in self-handicapping. Thus, we needed to first distinguish between various ways in which men and women might differ in their valuation of effort.

We generated an initial set of nearly 100 items which we narrowed to a smaller set of 35 items on the basis of pretesting. We (McCrea, Hirt, Milner, & Steele, 2006) then provided these items to a sample of 516 participants and conducted an exploratory factor analysis with oblimin rotation on these data. A scree plot revealed a clear break between the fourth and fifth factors and initial analyses revealed that a four factor solution was most interpretable. Specifically, we identified four scales that assessed qualitatively different aspects of valuing effort and the importance individuals attach to academic achievement (see Appendix). The Academic Concern scale is a seven-item measure that assesses an individual's concerns about performing well

academically. The second scale, called the Ability versus Effort Tradeoff scale, is a six-item scale that assesses the perceived importance of being seen by others as intelligent as opposed to a hard worker. This scale therefore assesses perceptions of the extent to which effort is viewed positively by others. The third scale is a five-item measure called the Prescriptive Effort Norm scale. This scale assesses the belief that people who put forth effort should be admired. Thus, this scale assesses endorsement of and adherence to a prescriptive effort norm. The final effort scale, an eight-item scale called the Worker scale, was designed to measure a more personal valuing of effort. Thus, the items on this scale assess the extent to which an individual sees him/herself as a hard-worker and personally values these characteristics. The Worker scale is different from many previously developed scales that measure societal norms or beliefs about the link between effort and success (e.g., Protestant Work Ethic, Mirels & Garrett, 1971). Rather, the scale reflects a view of self as someone who puts forth effort and personally values such effort. Each of the scales had adequate reliability and they were only weakly to moderately correlated ( $r_s < .37$ ).

After identifying these scales, we conducted an additional study focusing on the three effort scales (i.e., the Worker scale, Prescriptive Effort Norm scale, and Ability versus Effort Tradeoff scale) to validate the distinction made between them and to examine their relationships to theoretically relevant constructs. First, a confirmatory factor analysis examined whether these scales reflected a single or multiple underlying construct(s). The two models were tested using LISREL with maximum likelihood estimation on a sample of 1506 participants. A single factor solution provided a clearly inferior fit to the data ( $\chi^2 = 3228.96$ ,  $df = 151$ ,  $p < .001$ ,  $GFI = .82$ ,  $CFI = .86$ ,  $RMSEA = .12$ ) in comparison to the three factor solution ( $\chi^2 = 1344.08$ ,  $df = 148$ ,  $p < .001$ ,  $GFI = .91$ ,  $CFI = .93$ ,  $RMSEA = .07$ ). The scales again demonstrated adequate reliability and correlations between these measures were small to moderate in size ( $r_s < .43$ ).

Furthermore, both studies demonstrated gender differences across these three effort measures, such that women place greater value on effort than do men. However, the Worker scale appeared to us to be particularly likely to be related to the gender difference in behavioral self-handicapping for several reasons. First, we found that this scale predicted both Self-Handicapping Scale (SHS, Jones & Rhodewalt, 1982) scores ( $r = -.49$ ) as well as academic performance as measured by GPA ( $r = -.32$ ). Second, the worker scale was also moderately correlated with conceptually related variables such as Social Desirability ( $r = .27$ ; Crowne & Marlowe, 1960) and the Protestant Work Ethic ( $r = .23$ ; Mirels & Garret, 1971), and strongly correlated with Need for Achievement ( $r = .57$ ; Jackson, 1974) and the Big 5 Conscientiousness subscale ( $r = .64$ , John & Srivastava, 1999). However, the gender difference on the Worker scale was not reduced when controlling for these variables. Therefore, we (McCrea et al., 2006) suggested that gender differences on the Worker scale were likely to explain gender differences in the use of and reaction to behavioral self-handicapping.

The present research focuses on the behavioral self-handicap of effort withdrawal, and the extent to which the Worker scale serves as a mediator in the use of and reaction to individuals who engage in this form of behavioral self-handicapping. However, an important question that the astute reader may immediately consider is whether the Worker scale would explain only those behavioral self-handicaps dealing directly with a lack of adequate effort. We believe that the Worker scale is not restricted simply to explaining the use of the behavioral handicap of effort withdrawal and can account for gender differences in other types of behavioral handicaps as well. We found evidence that the Worker scale uniquely explained gender differences on those items of the SHS dealing with behavioral (e.g., “I would rather not take any drug that interfered with my ability to think clearly and do the right thing”), but not claimed (e.g., “I

suppose I feel ‘under the weather’ more often than most people”), forms of self-handicapping, even when controlling for related individual difference variables and the two other effort scales (McCrea et al., 2006). Moreover, the Worker scale partially mediated gender differences on SHS items mentioning taking drugs or getting sufficient sleep before a performance as well as inadequate expenditure of effort. Furthermore, two recent studies we and our colleagues (Flamm, 2007; McCrea & Flamm, 2006) have conducted examined the role of the Worker scale in the use of another form of behavioral self-handicapping, choosing to listen to distracting music during an intelligence test. We found correlations between music choice (assessed on a five-point scale ranging from very hindering to very helpful) and Worker scale scores in both studies:  $r(164) = -0.22, p < .01$  in the Flamm (2007) study, and  $r(43) = -0.35, p < .05$  in the McCrea and Flamm (2006) study. Thus, although beyond the scope of the present research, there is reason to believe that the Worker scale would explain gender differences in other types of behavioral handicaps, beyond lack of effort.

The present set of studies were specifically designed to test the role that effort beliefs play in accounting for gender differences in the use of reduced effort as a behavioral self-handicap as well as in reactions to reduced effort by others, and to rule out alternative explanations for these effects. We examined gender differences in observer reactions, in addition to actual behavior, for several reasons. First, this approach allows us to provide direct evidence supporting Hirt et al.’s (2003) explanation that women are more critical of others that behaviorally self-handicap because they value effort more than do men, which is an important goal in and of itself. Second, we believe that the way a given individual reacts to self-handicapping by someone else is an indirect assessment of the perceived benefits and costs of the handicap and therefore the likelihood that the individual would self-handicap in a similar manner. That is, if individuals do

not engage in a particular form of self-handicapping because they view such behavior negatively, they should evaluate others engaging in this behavior more negatively as well. If we can demonstrate that the Worker scale can account for gender differences observed in both of these paradigms, it attests to the importance that these effort beliefs play in determining one's own behavior as well as one's reaction to others.

### Study 1

In Study 1, we re-examined a set of data collected by McCrea and Hirt (2001). In that study, we found that men reported lower study effort than did women prior to a test in their introductory psychology course, even when controlling for variables such as self-esteem and past performance. Men went on to blame their subsequent poor performance on this handicap and thereby were able to protect self-esteem and conceptions of ability. Fortunately, as part of this study (but not reported by McCrea & Hirt, 2001), we also collected a number of additional measures germane to the present investigation.

First, we assessed participants' responses to the Worker scale. We predicted that the Worker scale would mediate these gender differences obtained in reported study effort.

Second, participants in this study also evaluated a target that withdrew study effort in a manner consistent with behavioral self-handicapping. Examining reactions to this scenario allowed us to test Hirt et al.'s (2003) explanation for the finding that women more negatively evaluate reduced effort than do men, as well as to provide additional evidence for our claim that women may not behaviorally self-handicap because they evaluate such behavior negatively. We again predicted that the Worker scale would be most effective in explaining these gender differences.

Third, we included a number of additional individual difference measures, allowing us to examine and (hopefully) rule out several alternative explanations for gender differences in reduced effort as a self-handicap. The first set of variables related to the importance of the task and feelings of uncertainty, as these variables have been shown to consistently moderate self-handicapping behavior (Harris & Snyder, 1986; Hirt et al., 2000; Self, 1990). It has been argued that women may not behaviorally self-handicap either because they place greater importance on academic performance or because they experience less self-doubt than do men. However, women actually tend to experience more self-doubt than do men (Oleson, Poehlmann, Yost, Lynch, & Arkin, 2000). Furthermore, it is difficult to explain why those high in self-doubt would be more critical of others who self-handicap.

Men could also be more motivated to behaviorally self-handicap than are women because they believe intelligence cannot be changed and that failure will indicate a lack of ability, which in turn has implications for the goals that individuals pursue in achievement contexts (Dweck & Leggett, 1988). Holding an entity theory should presumably make self-handicapping more likely, as the individual attempts to protect judgments of ability, rather than work toward mastery and self-improvement (Rhodewalt, 1994). Recent research (Elliot & Church, 2003; Rhodewalt, 1994) confirms the link between self-handicapping and the adoption of performance-avoidance goals, and Rhodewalt (1994) reports that women believe that effort improves ability more than do men. Therefore, it is possible that gender differences in self-handicapping originate with differences in implicit theories of intelligence.

We also examined several different aspects of valuing effort in addition to the Worker scale (McCrea et al., 2006). It could be that women fail to behaviorally self-handicap because they believe more strongly than do men that successful outcomes are best obtained by putting forth

effort. That is, women could have higher expectations in classic expectancy-value terminology (Mitchell, 1974; Vroom, 1964), perceiving that effort has greater utility in achievement.

Although we are not aware of any past research that has demonstrated such an effect of gender, it would be important to rule out this viable alternative explanation and illustrate that these gender differences persist even when controlling for possible gender differences in motivation.

Similarly, women may also simply enjoy putting forth effort more than men.

In summary, we predicted that gender differences in reports of reduced study effort as well as evaluations of someone else that studies inadequately would be best explained by more personal beliefs about the value of effort reflected by the Worker scale, over and above any differences resulting from other effort beliefs or motivation to self-handicap.

## Method

### *Participants and design*

Participants were 158 (68 male, 90 female) Introductory Psychology students at Indiana University-Bloomington. They received credit towards a course requirement for their participation. Participants completed questionnaires in three separate sessions. In the final session, they were randomly assigned to read one of two scenarios about a student preparing for an academic test. The outcome of the exam was manipulated ('A' or 'D'). Sixteen participants (six men, ten women) were eliminated from the study for failing to complete all three sessions.

### *Individual difference measures*

Participants completed the Worker scale (McCrea et al., 2006), responding to each statement using a scale ranging from 1 (strongly disagree) to 5 (strongly agree). In order to assess other aspects of valuing effort, we included the Beliefs in the Efficacy of Effort scale (BIEE, Poehlmann, 1994). The BIEE is a 55 item scale comprising three subscales: Effort leads



to/causes success (e.g., “The best predictor of how well I will do at something is how much effort I put into the task”), Ability leads to/causes success (e.g., “My level of ability alone determines how well I perform at something”), and Effort enjoyment (e.g., “I like engaging in tasks that require a lot of effort”). Participants indicated their agreement with these items on a 1 (disagree very much) to 6 (agree very much) scale.

A measure of implicit theories of intelligence was included (Dweck & Leggett, 1988). Three questions assessed the degree to which participants believed intelligence was a malleable rather than a fixed trait (e.g., “You have a certain degree of intelligence and you really can’t do much to change it”). Participants responded on a 1 (strongly agree) to 6 (strongly disagree) scale.

Participants completed the Academic Concern scale (McCrea et al., 2006). They responded to these items on a 1 (strongly agree) to 5 (strongly disagree) scale. Participants also completed the Subjective Overachievement Scale (SOS, Oleson et al., 2000), which is comprised of two subscales: Concern with performance and Self-doubt. Those high on both subscales are considered overachievers; those high in self-doubt but low in concern are considered likely to self-handicap. Participants were also administered the Rosenberg Self-Esteem Scale (SES, Rosenberg, 1965). Finally, participants completed the Self-Handicapping Scale (SHS, Jones & Rhodewalt, 1982). The SHS is designed to identify individuals likely to self-handicap and engage in other excuse-making behavior.

### *Self-handicapping scenario*

Participants read a scenario about Chris, a student who has an important upcoming test (cf. Hirt et al., 2003). Chris has not studied much and is feeling concerned about his performance. The night before the test, Chris is home studying for forty-five minutes before a friend stops by and invites him to see a movie. Chris and the friend go to the movie, and afterwards go out to

eat. Chris returns at 1am and goes right to bed. The outcome of Chris' test is then revealed to be either an 'A' or a 'D'.

### *Procedure*

The first session took place at the beginning of the semester in a mass testing format. Participants completed the Worker and Academic Concern scales. They then completed the SHS, self-esteem, and implicit theories of intelligence measures.

The second session took place approximately one week before the second exam in the introductory psychology course. Participants were told that the Department of Psychology was considering whether to change the format of the introductory psychology course and was soliciting feedback from students. Therefore, they were asked to provide information about the format of their course, characteristics of the instructor, and so on. Included among these items were the following questions: "How important is it for you to perform well on your P101 exam," "How fair do you find the exams in your P101 course?," "How satisfied are you with your performance in your P101 course so far?," "How interesting do you find the course material discussed in P101?." Participants answered these questions using a scale ranging from 1 (not at all) to 7 (very much). Participants were also asked to provide their score on their last exam.

Participants then completed several measures of self-reported study effort. Participants were asked to what extent they attended the course, took notes, and read the textbook, responding on 1 (not at all) to 7 (very much) scales. They indicated how much they planned to study and how much they had studied thus far (in hours). Lastly, they rated their overall study effort on a 1 (not at all) to 7 (very much) scale.

The final session took place approximately a week after the exam, after students had learned of their performance on the test. Participants first read the self-handicapping scenario, and

completed manipulation checks designed to ensure participants had carefully read the scenario. Participants then rated how sympathetic they felt towards Chris, how much they liked him overall, how similar they felt to Chris, and how well they could relate to Chris. Participants responded using 0 (not at all) to 6 (very much) scales. Participants were also asked how positively or negatively they felt towards Chris on a -6 (extremely negatively) to +6 (extremely positively) scale. Following the scenario questions, participants completed the SOS and the BIEE. They were then debriefed, thanked, and dismissed.

## Results

### *Overview*

To simplify the presentation of the results, a number of measures were combined into single indices. The measures of reported study effort prior to the exam were highly correlated and were therefore standardized and summed (reliability  $\alpha = .73$ ). Similarly, measures of evaluation of Chris were highly correlated and therefore standardized and summed (reliability  $\alpha = .81$ ). Dichotomous variables including gender (1 = men, 0 = women) were dummy coded, and all continuous variables were mean-centered (Aiken & West, 1991). The Aroian (Baron & Kenny, 1986; Goodman, 1960; Preacher & Leonardelli, 2001) version of the Sobel test was used to formally assess mediation.

### *Candidate mediators*

Correlations between the candidate mediators and reliabilities of these measures are presented in Table 1. Correlations were weak to moderate, indicating that none of the measures were tapping identical constructs.

Gender differences on the candidate mediators are presented in Table 2. There were reliable gender differences on the Worker scale, Academic Concern scale, SES, and the Self-doubt

subscale of the SOS.<sup>2</sup> As predicted, women reported valuing effort more on the Worker scale than did men, and also reported greater concern about performing well academically than did men. In addition, women had higher Self-doubt and lower self-esteem than did men. Given that those high in Self-doubt are more likely to handicap (Oleson et al., 2000), the gender difference on this scale is in the wrong direction to explain the gender difference. Such a difference would predict women self-handicapping more often than men. Nonetheless, we included Self-doubt as a possible mediator. There were no gender differences found on the measure of Implicit theories of intelligence, the SHS, the Concern subscale of the SOS, or any of the BIEE subscales.

#### *Reported study effort*

To demonstrate that gender differences in reported study effort were indeed reflective of self-handicapping, we conducted a regression analysis controlling for several motivational variables (see Table 3). To control for an expectancy-value model of motivation (Mitchell, 1974; Vroom, 1964), we included measures of the importance of the test and perceived fairness of the test. Lacking a measure of specific effort expectation, we included the more general Effort leads to success scale to assess the belief that effort was likely to improve performance. All interactions of these three variables were also calculated. We also controlled for reported performance on the last exam, interest in the course, and satisfaction with past performance in the course.<sup>3</sup> These variables were added along with gender in the first step of the model predicting reported study effort. In this initial model, the importance of the exam was a significant predictor of effort, such that individuals reported more study effort when they felt the test was more important. In addition, those that felt the course material was more interesting tended to report more study effort. The gender effect was also significant, indicating the men reported less study effort than did women. There were no other significant effects on reported study effort.

To test for mediation, we added the Worker scale, Academic Concern, self-esteem, and Self-doubt as a second step to this model. The addition of all mediators simultaneously allows for direct comparisons of the relative importance of each mediator (see also Kenny, Kashy, & Bolger, 1998). Only the Worker scale was a significant predictor, indicating those valuing effort more reported more study effort. Adding these terms to the model significantly increased the amount of variance explained and rendered the gender effect on reported study effort non-significant. The Worker scale reliably mediated this gender effect (Sobel test  $z = 2.61, p < .01$ ).

In an additional analysis, we also tested the unique effects of the Worker scale. The Worker scale was again a significant predictor of reported study effort ( $\beta = 0.379, t = 4.02, p < .001$ ), and added significantly to the amount of variance explained,  $R^2 = 0.31, \Delta R^2 = 0.10, F_{\text{Change}}(1,113) = 16.17, p < .001$ . Furthermore, the inclusion of the Worker scale eliminated the gender effect ( $\beta = -0.110, t = 1.30, p = .20$ ), reliably mediating this effect (Sobel test  $z = 2.49, p < .05$ ).

### *Evaluations of Chris*

Participants correctly answered on average 8.87 out of 9 manipulation check items, indicating that they paid careful attention to the details of the passage. Initial analyses revealed no effects of grade for evaluations of Chris (all  $F$ s  $< 1$ , ns) and so this variable was dropped from the analyses. The model predicting evaluations of Chris was the same as that used to predict reported study effort. Of the covariates, only importance placed on the exam had an effect, such that those placing more importance on their own exam more negatively evaluated Chris. Gender also significantly predicted evaluations. Consistent with the results of Hirt et al. (2003), men rated Chris more positively than did women.

In order to test for mediation, the Worker scale, Academic Concern, self-esteem, and Self-doubt were again added as a third step to the model predicting evaluations of Chris (see Table 4).

Of the potential mediators, only the Worker scale was a significant predictor of evaluations, indicating that those valuing effort more evaluated Chris more negatively. Addition of the mediators to the model significantly increased the amount of explained variance and reduced the gender effect on evaluations of Chris. The reduction in the gender effect due to the Worker scale was reliable (Sobel test  $z = 2.07$ ,  $p < .05$ ), demonstrating partial mediation.

In an additional analysis, we also tested the unique effects of the Worker scale. The Worker scale was again a significant predictor of evaluations of Chris ( $\beta = -0.313$ ,  $t = 3.39$ ,  $p < .01$ ), and added significantly to the amount of variance explained,  $R^2 = 0.28$ ,  $\Delta R^2 = 0.07$ ,  $F_{\text{Change}}(1,115) = 11.49$ ,  $p < .01$ . Furthermore, the inclusion of the Worker scale eliminated the gender effect ( $\beta = 0.158$ ,  $t = 1.82$ ,  $p = .07$ ), reliably mediating this effect (Sobel test  $z = 2.30$ ,  $p < .05$ ).

### Discussion

The results of Study 1 demonstrate that the Worker scale explains the gender difference found in reports of study effort, as well as the gender difference in reactions to a target that withdraws effort. These robust gender differences were not explained by related constructs such as Academic Concern, the belief that Effort leads to success, or self-esteem.

It is significant that the Worker scale mediated both reports of participants' own effort behavior as well as evaluations of a target. Individuals scoring higher on the scale appear to view a lack of effort as unacceptable. These findings provide direct support for the explanation of gender differences in observer reactions to self-handicapping provided by Hirt et al. (2003). That is, if an individual chooses to self-handicap by withdrawing effort, women are likely to be more critical of this behavior than are men, precisely because effort is personally important for them and, as a result, they view inadequate effort more suspiciously. These results also provide indirect support for the notion that women do not self-handicap by reducing effort because they

view such behavior negatively. Evidence for this claim is provided by the significant negative correlation found between evaluations of Chris and reported study effort ( $r = -.38, p < .001$ ).

We controlled for a number of motivational variables to ensure that any differences in reported study effort were likely to reflect self-handicapping. Indeed, gender differences were found despite controlling for the expected value of effort (Mitchell, 1974; Vroom, 1964), past performance, and interest in the course. Furthermore, McCrea and Hirt (2001) reported that these reports of reduced effort were correlated with actual performance on the course exam and predicted higher attributions to (lack of) effort and higher self-esteem following poor performance for men. Thus, in our view, these gender differences in reported study effort appeared to serve as a self-handicap, and the Worker scale alone explained these differences. Nonetheless, due to the self-reported nature of these data, we cannot yet definitively answer whether gender differences in actual behavioral self-handicapping are explained by the Worker scale. We return to this issue in Study 3.

## Study 2

Buoyed by these suggestive initial findings, we next sought to further delineate the nature of the gender difference in valuing effort. As we have discussed, the Worker scale measures the more personal value placed on effort, i.e., the extent to which a person views him/herself as someone that consistently puts forth effort. But can we be sure that it is this facet of valuing effort that is driving the observed results? Couldn't it be the case that gender differences in perceived societal norms concerning the importance of putting forth effort might play a role here? Given the evidence supporting the impression management aspect of self-handicapping (Hirt et al., 2003; Kolditz & Arkin, 1982), whether men and women believe putting forth effort is normative could be of crucial importance to whether they use a lack of effort as a handicap and

how they view someone that behaves in this manner. Therefore, in Study 2, we included the Prescriptive Effort Norms and Ability versus Effort Tradeoff scales developed by McCrea et al. (2006) to assess these more normative beliefs about effort. As described earlier, the Prescriptive Effort Norms scale assesses perceived societal norms encouraging the use of effort, such that individuals should always put forth their best effort. The Ability versus Effort Tradeoff scale, on the other hand, assesses the belief that society values the demonstration of ability to a greater extent than the display of effort. Presumably, those who believe society values effort are more likely to be critical of those who withdraw effort. Furthermore, McCrea et al. (2006) found that consistent gender differences emerge on these two scales. Thus, in Study 2, we examined whether a consideration of these alternative aspects of the valuation of effort would in any way qualify our finding that the Worker scale mediates gender differences in observer reactions to lack of effort. Although the normative beliefs measured by the Ability versus Effort Tradeoff and Prescriptive Effort Norms scales may also underlie these reactions, we predicted that the Worker scale, because it measures more personal value placed on effort, would play a more significant role in how individuals react to a lack of effort by another.

A second goal of Study 2 was to examine the robustness of gender differences in observer reactions to a target that puts forth inadequate effort. Playing devil's advocate for a moment, one could interpret the results of Study 1 as simply demonstrating that people who state that they personally value effort do indeed put forth effort and expect others to do the same. Thus, we wanted to ask whether our results reflect more general negative reactions to any form of reduced effort or a more specific aversion to instances of behavioral self-handicapping. Therefore, we examined whether there exist any extenuating circumstances that might cause women to be less critical of a lack of effort.



In initial research on this question, Hirt et al. (2003) varied whether the target had done any prior preparation and whether he initiated going to the movies or not. These manipulations were intended to make available alternative motives (i.e., feeling prepared, succumbing to peer pressure) to the conclusion that the target was self-handicapping. Men indeed were more likely to give the target the “benefit of the doubt” when the target had done some prior preparation or the trip to the movies was initiated by a friend, but women continued to be critical of the target. Of course, going to the movies rather than studying before an important test is a relatively indefensible behavior. Study 2 was designed to provide a stronger test of the hypothesis that certain forms of low effort would be viewed as acceptable by women as well as men.

To examine this question, two characteristics of the target’s behavior were varied in Study 2. Specifically, the behavior of the target varied in intentionality as well as social desirability. Past research by Baumgardner and Levy (1988) found that when effort withdrawal was unintentional, targets were viewed more favorably than when it was intentional. Hence, we might expect that women as well as men might be less critical of a target who fails to study through no fault of their own (e.g., accident, power failure). Thus, we manipulated intentionality by including a condition in which the target intended to study, but got back late due to car trouble (unintentional) along with one in which the target chose to go out all night (intentional).

The effects of the social desirability of the target’s behavior have not yet been examined in the context of self-handicapping. A lack of effort in the service of another socially desirable behavior (e.g., helping a friend), by fulfilling other important social norms and obligations, may override the personal value placed on effort or norms concerning the importance of working hard. To test this notion, we introduced an additional condition in Study 2 in which the target foregoes studying in order to help a friend. Reactions to this more socially desirable behavior

were then compared to perceptions of the rather indefensible behavior of going to the movies. We expected that this manipulation would affect women's reactions to the target such that they would be as sympathetic as would men to the target when he engages in a more socially desirable behavior. In this case, women should be less likely to view the behavior as reflecting self-handicapping or laziness, and therefore should evaluate the target more positively.

## Method

### *Participants*

Participants were 397 introductory psychology students (185 men, 212 women) at Indiana University–Bloomington. They participated in this study in partial fulfillment of a course requirement. Participants were tested in groups ranging in size from 1 to 25.

### *Materials*

The passage used for this study was similar to that used in Study 1. Participants again read about the actions of a character named Chris who puts forth inadequate effort prior to an exam. The social desirability of Chris' behavior, the intentionality of the action, and Chris's exam grade (A or D) were factorially manipulated, creating eight conditions.

In the socially desirable condition, Chris offered to drive his friend George to the airport, so that George may fly home to visit a sick relative. George tells Chris that he can get a ride from a different friend, but Chris insists on taking George himself. Thus, Chris could have easily chosen to study, but does not. In the socially undesirable condition, Chris invites George to see a movie, as in Study 1.

Participants in the intentional condition read that Chris stayed out all evening (as in Study 1). Participants in the unintentional condition read that Chris experienced car problems on his drive

home that kept him out later than anticipated. This manipulation paralleled that used by Baumgardner and Levy (1988).

### *Procedure*

Participants were randomly assigned to one of the eight conditions. The procedure was adapted from Study 1. Participants were told that they would be reading a description of a character for a book, and were asked to give their impressions of him. Participants read the vignette and completed the questions that followed. Dependent measures included an overall evaluation of Chris, reported sympathy for Chris, reported liking of Chris, perceived similarity to Chris, and ability to relate to Chris. Participants responded using 0 (not at all) to 6 (very much) scales, with the exception of the overall evaluation which was made on a -6 (extremely negative) to +6 (extremely positive) scale. They also answered a few manipulation check items to ensure they had paid close attention to the scenario. All participants then completed the Worker scale, the Prescriptive Effort Norms scale, and the Ability versus Effort Tradeoff scale.

## Results

### *Overview*

The dependent measures were again combined by standardizing and summing the items to form an overall evaluation index (reliability  $\alpha=.83$ ). Preliminary analyses revealed no effects of intentionality and only a main effect of outcome indicating that successful targets were evaluated more positively. To simplify the presentation of the results, these terms were dropped from further analysis. For the regression analyses, dichotomous variables were dummy coded (gender: 1 = men, 0 = women; type of behavior: 1 = desirable, 0 = undesirable) due to the inclusion of interaction terms with continuous variables, and all continuous variables were mean-centered

(cf., Aiken & West, 1991). For significant interactions, predicted scores were plotted at +1 SD and -1 SD of continuous variables.

Reliability for the Worker scale was again good ( $\alpha=.80$ ), but somewhat low for the Ability vs. Effort Tradeoff ( $\alpha=.60$ ) and the Prescriptive Effort Norms scales ( $\alpha=.61$ ). The Worker scale was correlated  $r = -.27, p < .001$  with the Ability vs. Effort Tradeoff scale and  $r = .45, p < .001$  with the Prescriptive Effort Norms scale. The correlation between the Ability vs. Effort Tradeoff scale and Prescriptive Effort Norms scale was  $r = -.39, p < .001$ . Thus, although related, the effort scales appeared to be assessing different constructs. There were also no differences by experimental condition ( $F_s < 2.03, ns$ ) on any of the individual difference measures, indicating that random assignment had been successful. Participants scored on average 8.9 out of 9 correct on the manipulation check items, again indicating they had paid careful attention to the details of the passage.

#### *Gender differences on evaluation*

A 2 (Gender) x 2 (Type of handicap) ANOVA was conducted on the evaluation of Chris. Means are presented in Figure 1. This analysis revealed a main effect of gender, such that men reported a more positive evaluation of Chris than did women,  $F(1,312) = 4.68, p < .05, \eta^2 = .01$ . This effect was qualified by a significant Gender x Type of behavior interaction,  $F(1,312) = 4.73, p < .05, \eta^2 = .01$ . Women more negatively evaluated the undesirable behavior than did men,  $t(156) = 3.06, p < .01, \eta^2 = .07$ , paralleling the effects observed in Study 1. However, men and women evaluated the socially desirable behavior equivalently,  $t(156) < 1, ns, \eta^2 = .00$ .

#### *Mediation analyses*

As an initial analysis, a regression analysis predicting evaluation was conducted with gender, type of behavior, and their interaction entered as a first step in the model, and the Worker scale,

Ability vs. Effort Tradeoff scale, and Prescriptive Effort Norms scale added as a second step.

Inclusion of the terms in the second step added significantly to the overall model,  $R^2 = .15$ ,  $\Delta R^2 = .11$ ,  $F_{\text{Change}}(6,301) = 6.37$ ,  $p < .001$ . The Ability vs. Effort Tradeoff scale and Prescriptive Effort Norms scale had no relationship to evaluation of Chris ( $t_s < 1.61$ , ns). However, the Worker scale ( $\beta = -0.446$ ,  $t = 5.18$ ,  $p < .001$ ) and the interaction of Worker scale x Type of behavior ( $\beta = 0.190$ ,  $t = 2.25$ ,  $p < .05$ ) were both found to be significant predictors of evaluations of Chris while controlling for gender. The Worker scale x Type of behavior interaction closely resembled the Gender x Type of behavior interaction. Predicted scores are presented in Figure 2.

Individuals high on the scale, much like women, showed more disapproval of Chris when he went to the movies compared to those low on the scale (simple-slope =  $-.268$ ,  $t = -5.96$ ,  $p < .001$ ). Individuals high and low on the scale equivalently evaluated the target who gives his friend a ride to the airport (simple-slope =  $-.049$ ,  $t = -1.03$ , ns).

Mediational analyses were then conducted within the undesirable handicap condition, in which gender differences were found on the evaluation measure. Gender differences on the candidate mediators within this condition are presented in Table 5. Replicating Study 1, women reported personally valuing effort more on the Worker scale than did men. As expected, women also reported placing more importance on effort on the Prescriptive Effort Norms scale compared to men. No differences were observed on the Ability vs. Effort Tradeoff scale in this condition. We therefore did not include this term in the mediation analyses.

To test for mediation, gender was included as a first step in the regression model predicting evaluations of Chris in the undesirable handicap condition, and the Worker scale and Prescriptive Effort Norms scale were added as a second step (see Table 6). The Worker scale was a significant predictor of evaluations of Chris, whereas the Prescriptive Effort Norms scale was

not. The inclusion of these terms added to the explained variance. The gender effect was also significantly reduced (Sobel test  $z = 2.82$ ,  $p < .01$ ), indicating the Worker scale partially mediated gender differences in evaluation of Chris when he went to the movies.

In an additional analysis, we also tested the unique effects of the Worker scale. The Worker scale was again a significant predictor of evaluations of Chris in this model ( $\beta = -0.445$ ,  $t = 6.19$ ,  $p < .001$ ), and added significantly to the amount of variance explained,  $R^2 = 0.25$ ,  $\Delta R^2 = 0.19$ ,  $F_{\text{Change}}(1,155) = 38.35$ ,  $p < .001$ . Furthermore, the inclusion of the Worker reduced the gender effect ( $\beta = 0.145$ ,  $t = 2.02$ ,  $p < .05$ ), partially mediating this effect (Sobel test  $z = 2.87$ ,  $p < .01$ ).

### Discussion

Study 2 replicated past work showing that women more negatively evaluate individuals that withdraw effort in a manner consistent with self-handicapping (Hirt et al., 2003). However, we also found that the social desirability of the behavior moderates this effect. Women more positively evaluated a target who withdrew effort in order to help a friend, even if such behavior was unnecessary. Contrary to our predictions, there were no effects of the intentionality manipulation. It is possible that this manipulation was not powerful enough; many women may have felt that the initial choice to delay studying and withdraw effort, even if one intended to do it later, was inexcusable. Thus, in a future study, it may be interesting to include a condition in which the reduced effort was due to a truly unforeseen event. Nonetheless, Study 2 illustrated that the social desirability of a behavior that can serve as a self-handicap is a critical variable, particularly for women. Women are likely to be more forgiving of a self-handicapper when the behavior satisfies other socially prescribed norms, such as helping others. However, when the behavior is not justified by such selfless action, women are less tolerant and expect adequate effort. These findings support the conclusions of Hirt et al. (2003) that the suspected motives of

the target determine responses to self-handicapping behavior. These judgments are nonetheless sensitive to situational cues, suggesting that women are more critical of a lack of effort that is likely due to self-handicapping or laziness. Only when the lack of effort could be considered as fulfilling some other more powerful social norm do women evaluate this behavior as positively as do men. These findings also support our contention that women may not self-handicap themselves by withdrawing effort because they view such behavior negatively.

Importantly, we replicated the gender difference in beliefs concerning effort (see also McCrea et al., 2006). Not only are there gender differences in the Worker scale, but also the Prescriptive Effort Norms scale. Thus, it seems that men and women differ not only in the personal value they place on effort (indexed by the Worker scale), but also in their beliefs of how much effort individuals that put forth effort should be admired (indexed by the Prescriptive Effort Norms scale). However, gender differences in observer reactions to the target were mediated only by Worker scores and not by the other effort scales. These effects even paralleled the interaction pattern with type of behavior described above. Thus, it appears that more personal valuations of effort, but not more normative beliefs, play an important role in determining these gender differences. These findings again support our view that women may not behaviorally self-handicap because they more generally view such behavior negatively.

### Study 3

Thus far, we have demonstrated that the Worker scale mediates gender differences in both audience reactions to reduced effort and reports of study effort prior to an exam. However, we have not yet examined whether these beliefs underlie gender differences in observed behavioral self-handicapping, nor have we examined the role of the Prescriptive Effort Norms or Ability vs. Effort Tradeoff scales in behavioral self-handicapping. Even though the reports of study effort in

Study 1 were significantly correlated with subsequent performance (see also McCrea & Hirt, 2001), the authenticity of these reports as reflecting actual behavior can still be questioned. In addition, a skeptic could argue that the Worker scale is simply a more general measure of the same type of reporting of degree of effort. Therefore, it is crucial to demonstrate that the Worker scale also mediates gender differences in observed effort prior to an exam. Study 3 was therefore designed to replicate past studies demonstrating gender differences in behavioral self-handicapping (e.g., Berglas & Jones, 1978; Hirt et al., 2000).

The setting, tasks, and instructions for Study 3 were modeled after Hirt et al. (2000). Participants were given noncontingent success feedback, as this manipulation has been repeatedly shown to lead men to behaviorally self-handicap (Berglas & Jones, 1978; Hirt et al., 2000; Kolditz & Arkin, 1982). Thus, Study 3 was designed to create the kind of uncertainty and evaluative concern among participants shown to promote the use of behavioral self-handicapping. As in past studies (Hirt et al., 1991; Hirt et al., 2000), we manipulated the availability of a lack of practice effort as a handicap. Half of the participants in the study were assigned to a “practice matters” condition, in which they were told that the exam was only valid with prior practice, whereas the other half were assigned to a “practice does not matter” condition, in which they were told practice had no real effect on performance. Thus, lack of effort could serve as a self-handicap only when practice matters, and the practice doesn’t matter condition provides a baseline assessment of motivation to practice. If men are just “lazier” or do not care about doing well on the test, we should find only main effects of gender on practice effort. Rather, we predicted that men would practice less than women, but only when told that practice was required for the exam to be valid. Such a pattern would indicate self-handicapping



on the part of men (see also Hirt et al., 1991; Hirt et al., 2000). We also included measures of the importance of the test and beliefs in the usefulness of practice to bolster this interpretation.<sup>4</sup>

We then examined whether any gender differences in practice effort would be explained by effort beliefs. We therefore included the Prescriptive Effort Norms, Ability vs. Effort Tradeoff, and Worker scales in Study 3. We expected the Worker scale would best explain any gender differences in practice effort.

## Method

### *Participants and design*

Participants were 188 (92 male and 96 female) introductory psychology students at Indiana University–Bloomington. They received research participation credit in return for being involved in the study. Participants were randomly assigned to either a practice matters or practice does not matter condition (see also Hirt et al., 1991; Hirt et al., 2000). The study was therefore a single-factor design, with gender allowed to vary.

### *Materials and procedure*

Participants first completed the Worker, Ability vs. Effort Tradeoff, and Prescriptive Effort Norms scales. They were told these measures were part of an unrelated study. Participants were then informed that the experiment involved comparing a newer, culture-fair nonverbal intelligence test to a more traditional verbal intelligence test. The experimenters explained that the nonverbal test, called the CFIT (Cattell & Cattell, 1961), eliminated cultural bias and was a good predictor of academic and career success (see Hirt et al., 2000).

Participants first completed a 15-item multiple-choice analogy test on a computer. The test served as a noncontingent success manipulation (cf., Berglas & Jones, 1978; Hirt et al., 2000). Participants were told this test was a traditional verbal intelligence test. The items chosen were

particularly difficult. After completing the exam, the computer displayed a message informing participants that they had scored 12 out of 15 items correct, and that this score placed them in the top 10% of college students. They were then told that they would be expected to score highly on the nonverbal exam as well, and that the researchers would compare their two scores to see if they matched this high expectation. These instructions were given in order to make participants feel uncertain of their ability to perform well on the exam, as well as to heighten the impression that the experimenter would be closely observing their performance (see also Hirt et al., 2000). Following these instructions, participants completed an eight-item measure of the importance of the exam (e.g., “It is important to me that I do well on this test,” see also Hirt et al., 2000) using a 1 (not at all) to 7 (very much) scale.

Participants then were assigned to one of two practice instruction conditions (see Hirt et al., 1991; Hirt et al., 2000). Those in the practice matters condition were told that past research had shown that scores on the nonverbal exam were not diagnostic unless individuals had some prior experience with the items, and that individuals without such practice would score lower than their actual intelligence would warrant. Those in the practice does not matter condition were told that past research had shown that practice had no effect on scores on the exam. Participants in both conditions were then told that they would be given a chance to practice prior to the exam.

Participants were then given a practice booklet containing 18 practice items. Participants were told that they could practice for as long or as little as they liked, and could start the exam whenever they felt ready. The time participants spent practicing was surreptitiously recorded, along with the number of practice items completed. They then completed the 15 item exam. Following the exam, participants completed a manipulation check, indicating to what extent they believed that practice helped on the exam using a 1 (not at all) to 7 (very much) scale. They also

indicated whether they had felt pressured by the experimenter to practice, and whether they believed the test was a valid measure of intelligence. These items were also assessed on a 1 (not at all) to 7 (very much) scale.

## Results

Time spent practicing ( $M = 168.56$  seconds,  $SD = 131.27$ ) and number of practice items completed ( $M = 7.54$ ,  $SD = 6.36$ ) were highly correlated ( $r = .67$ ) and were therefore standardized and summed to create a single index of practice effort (reliability  $\alpha = .80$ ), as in prior research (e.g., Hirt et al., 2000). The reliability for the candidate mediators ranged from marginally acceptable (Ability vs. Effort Tradeoff  $\alpha = .66$ ; Prescriptive Effort Norms  $\alpha = .60$ ) to good (Worker  $\alpha = .81$ ). The Worker scale ( $M = 24.98$ ,  $SD = 5.92$ ) was correlated  $r = -.25$ ,  $p < .001$ , with the Ability vs. Effort Tradeoff scale ( $M = 13.32$ ,  $SD = 4.59$ ) and  $r = .46$ ,  $p < .001$ , with the Prescriptive Effort Norms scale ( $M = 17.10$ ,  $SD = 2.33$ ). The correlation between the Ability vs. Effort Tradeoff scale and Prescriptive Effort Norms scale was  $r = -.29$ ,  $p < .001$ .

### *Manipulation check*

An ANOVA on the manipulation check assessing participant's beliefs about practice revealed a significant effect of the instruction manipulation,  $F(1,184) = 9.89$ ,  $p < .01$ , such that those in the practice matters condition believed that practicing was more important ( $M = 3.84$ ) than did those in the practice does not matter condition ( $M = 2.99$ ). Thus, it appears that participants properly understood and believed the practice instructions. Importantly, there were no effects of gender on this measure ( $F_s < 1$ , ns), indicating that men believed that practice was just as important to performing well on the CFIT as did women.

### *Practice effort*

To provide additional evidence that reduced practice effort in the practice matters condition reflected behavioral self-handicapping, we controlled for reports of the importance of the exam ( $\alpha = .86$ ), the perceived validity of the exam, and perceived pressure to practice as covariates.<sup>5</sup> Therefore, a 2 (gender) x 2 (practice instruction) ANCOVA, including these three variables as covariates, was conducted on the practice effort index. Adjusted means, after controlling for the three covariates, are presented in Figure 3. There were no effects of the covariates (all  $F$ s  $< 1.99$ ,  $p$ s  $> .16$ ). The effect of practice instruction approached significance,  $F(1,181) = 2.90$ ,  $p = .09$ ,  $\eta^2 = 0.01$ . There was a significant effect of gender,  $F(1,181) = 11.79$ ,  $p < .01$ ,  $\eta^2 = 0.06$ , such that men practiced less than women.

In addition, a significant interaction of Practice instruction x Gender was observed,  $F(1,181) = 4.61$ ,  $p < .05$ ,  $\eta^2 = 0.02$ . This interaction revealed that men practiced less than women in the practice matters condition,  $F(1,95) = 16.36$ ,  $p < .001$ ,  $\eta^2 = 0.13$ , but not in the practice does not matter condition,  $F(1,83) = 1.70$ ,  $p = .20$ ,  $\eta^2 = 0.02$ . Put differently, whereas women practiced significantly more when told practice mattered,  $F(1,91) = 6.75$ ,  $p < .05$ ,  $\eta^2 = 0.07$ , men practiced non-significantly less when told practice mattered,  $F(1,87) < 1$ , ns,  $\eta^2 = 0.00$ . This pattern replicates findings obtained in Hirt et al. (2000) and again demonstrates that men self-handicapped in the practice matters condition by not putting forth adequate practice effort.

#### *Mediational analyses*

Next, we sought to test mediation of the significant gender effect on the practice index within the practice matters condition. In this condition, gender differences were observed on both the Worker and Prescriptive Effort Norms scales (see Table 7). Regression analyses were then used to test for mediation within the practice matters condition. Dummy coding was used for the gender variable (Men = 1, Women = 0). To ensure any observed effects were not due to the

importance of the exam, perceived validity of the exam, or perceived pressure to practice, we conducted our analyses on residual practice scores after removing the effects of these variables. There were no gender differences on the Ability vs. Effort Tradeoff scale and preliminary analyses revealed this variable did not predict the practice index, and so this variable was not included in the analysis. Therefore, gender was added as a first step in the analysis, and the Worker scale and Prescriptive Effort Norms scale were added as a second step. The results of these mediational analyses are summarized in Table 8. The initial model including only gender revealed a significant gender effect. In the mediation model, the Worker scale predicted practice index residual scores while controlling for gender differences, whereas the Prescriptive Effort Norms scale did not. The addition of these terms added significantly to the model. The inclusion of the Worker scale reduced the gender effect (Sobel test  $z = 2.44, p < .05$ ), although not completely eliminating it. Thus, the Worker scale partially mediated gender differences in the practice matters condition, whereas the Prescriptive Effort Norms scale did not.

In an additional analysis, we also tested the unique effects of the Worker scale. The Worker scale was again a significant predictor of practice effort ( $\beta = 0.439, t = 4.99, p < .001$ ), and added significantly to the amount of variance explained,  $R^2 = 0.31, \Delta R^2 = 0.18, F_{\text{Change}}(1,97) = 24.86, p < .001$ . Furthermore, the inclusion of the Worker scale reduced the gender effect ( $\beta = -0.241, t = 2.73, p < .01$ ), partially mediating this effect (Sobel test  $z = 2.57, p < .05$ ).

### Discussion

Study 3 nicely replicated the findings of Hirt et al. (2000), illustrating that men practiced less than did women prior to an important intelligent test, but only when told that a lack of practice effort rendered the CFIT less valid. It is perhaps interesting to note that these effects are not simply overall mean differences, but reflect the tendency of individuals to completely fail to take

advantage of a supposedly useful practice opportunity. Indeed, within the practice matters condition, fourteen men (27%) and only three women (6%) did not complete a single practice item, which in itself was a significant difference. Similar numbers are observed when considering time spent practicing. Thirteen men (25%) and only two women (4%) spent less than 30 seconds on the practice exam. Furthermore, these effects were not due to differences in the perceived usefulness of practice, importance placed on the exam, perceived validity of the test, or perceived pressure to practice. Taken together, these results suggest reduced practice effort by men was indicative of behavioral self-handicapping. If men were simply lazier than women, we would expect to find only a main effect of gender rather than an interaction with practice instruction. The results clearly demonstrate a gender difference was only observed in the practice matters condition, in which lack of effort could serve as a self-handicap and provide an excuse for failure.

As predicted, we found evidence that the Worker scale, and not normative beliefs about effort reflected by the Prescriptive Effort Norms or Ability vs. Effort Tradeoff scales, mediated these differences in behavioral self-handicapping. These findings also support the results of Study 1, demonstrating that the Worker scale mediated gender differences in the reports of study effort. Thus, there is consistent evidence that an increased personal value placed on effort reduces the appeal of low actual (Study 3) as well as reported effort (Study 1) as a form of self-handicapping. That the gender difference remains significant after controlling for the Worker scale indicates that there may be additional mediators of this effect, but these beliefs do appear to be an important cause of gender differences in this behavior.

## General Discussion

For years, the self-handicapping literature has lacked a clear explanation for the robust gender differences that exist in the use of behavioral self-handicapping. Furthermore, our own recent work (Hirt et al., 2003) demonstrated that women not only fail to use lack of effort as a self-handicap, but also respond more negatively to targets who engage in such actions. Our primary goal in the present research was to investigate whether gender differences in the value placed on effort might account for these robust gender differences. To investigate this question, we examined the roles that several different aspects of effort valuation might play in producing these effects. The present set of studies provide strong support for the notion that gender differences in the extent to which individuals reduce effort as a self-handicap and in the reaction to others who do so are explained by differences in the personal value placed on effort. Women place more personal importance on putting forth effort, as reflected by scores on the Worker scale, than do men. As a result, women are less likely than men to engage in self-handicapping behaviors that are inconsistent with this belief, and evaluate such behavior on the part of others more negatively than do men.

The fact that scores on the Worker scale can account for not only the gender differences in the use of effort as a handicap, but also reactions to others' use of this behavior, attests to the power of these beliefs. The Worker scale explains both women's negative reactions to the behavioral self-handicapping of others as well as their own reluctance to engage in this behavior. We view this result as further support for our argument that it is the characteristics of the self-handicapping behavior that lead to gender differences in their use. Whereas gender differences are observed in the use of behavioral handicaps, women appear just as likely or even more likely than men to engage in other forms of self-handicapping, such as claiming stress or bad mood (Hirt et al., 1991; Koch et al., 2003; Rhodewalt, 1990). We have argued previously (Hirt et al.,

2003) that women appear to view behavioral self-handicaps as unacceptable and perhaps are therefore less likely to engage in them themselves. The present results supported this view, demonstrating that women are more critical of insufficient effort, particularly when it can more easily be characterized as self-handicapping rather than fulfilling some other important social norm (such as helping a friend). Furthermore, evaluations of such behavior were correlated with reports of reduced study effort. Lastly, the present work finds that the effort beliefs indexed by the Worker scale account for gender differences in these evaluations as well as the use of reduced effort as a self-handicap. Thus, one must consider the appeal of the handicap itself when considering the likelihood that certain individuals will self-handicap in that manner. Certainly, the beliefs reflected by the Worker scale should not preclude the same individual from self-handicapping by some other means. Indeed, we have recently found that individuals scoring high in the Worker scale are just as likely to claim increased stress when it serves as a viable self-handicap as those scoring low on this scale (Hendrix & Hirt, 2006), suggesting that they are just as likely to self-handicap when provided with a more appealing potential handicap.

The present studies not only provide the first viable explanation of which we are aware that account for these robust gender differences, but also allow us to rule out a number of alternative explanations for these findings. Although we found gender differences on variables related to the motivation to self-handicap, such as Academic Concern, the Self-doubt subscale of the SOS, and self-esteem, none of these variables explained gender differences in effort or reactions to a target that withdraws effort. We also examined a number of different aspects of importance placed on effort. However, there did not appear to be consistent gender differences on Beliefs in the Efficacy of Effort (Poehlmann, 1994) or Dweck's (Dweck & Leggett, 1988) implicit theories of intelligence. Replicating our past work (McCrea et al., 2006), we did observe that women



perceive that effort is more normative on the Prescriptive Effort Norms scale. However, only the more personal value placed on effort assessed by the Worker scale, not these other dimensions of effort valuation, mediated observed gender differences. Finally, the effects of the Worker scale on reported effort in Study 1 and on actual effort in Study 3 were found even controlling for expected-value and task importance, suggesting it is not differences in desire to perform well per se that explains the gender differences in behavioral self-handicapping. Thus, the present research has demonstrated that the aspects of effort valuation measured by the Worker scale uniquely account for these reactions to behavioral handicapping, over and above these other predictors.

*Generality of the Worker scale as an explanation for gender differences in behavioral self-handicapping*

One limitation of the present work is that we have focused exclusively on the behavioral handicap of reduced effort. It is therefore unclear whether the Worker scale would also account for gender differences in the use of other types of behavioral self-handicaps, such as alcohol or drug use. It might be necessary to expand the scope of the Worker scale to include other handicap-specific self-views (e.g., seeing oneself as someone who does not drink to excess). However, to the extent that the Worker scale taps a desire to put forth one's best effort, we might expect similar results for other types of behavioral handicaps. We have argued here and elsewhere (Hirt et al., 1991; Hirt et al., 2003) that it is differences in the perceived benefits and costs of self-handicaps that determines their use. Thus, individuals with a strong aversion to behaving in a lazy or unmotivated fashion are likely to have a similar distaste for other forms of behavioral self-handicapping. Thus, there is reason to believe that the Worker scale might explain gender differences in other types of behavioral handicaps. Indeed, we mentioned in the

Introduction that in our work (McCrea et al., 2006) developing the Worker scale, we found that the Worker scale appears to at least partially mediate gender differences on SHS items that deal with behavioral handicaps unrelated to effort. Furthermore, the Worker scale is correlated with the selection of distracting music as a self-handicap (Flamm, 2007; McCrea & Flamm, 2006). Thus, while it remains for future research to examine in more detail whether the Worker scale can explain gender differences in other types of behavioral self-handicapping, we firmly believe that its purview extends beyond the handicap of effort withdrawal to behavioral self-handicapping more broadly.

A somewhat related issue concerns whether the Worker scale would explain differences in self-handicapping in other domains beyond just intelligence and academics. Although several of the items on this scale are limited to academic situations, others concern the role of effort in performance more generally (e.g., “I work hard to be successful at whatever I do”). Furthermore, the arguments we have made regarding the ability of the scale to predict other types of handicaps suggests it would likewise predict the same types of behavior in other domains. Clearly, these are important questions which require further study.

*Do Worker scale scores reflect a general adherence to norms?*

The implications of this work for gender differences in other self-defeating behaviors is also worthy of further research. Men are more likely than women to engage in a number of risky, self-destructive, and antisocial behaviors, such as drug and alcohol use, aggression, and reduced medical compliance (Baumeister & Scher, 1988; Courtenay, McCreary, & Merighi, 2002; Eagly & Steffen, 1986; Substance abuse and Mental Health Services Administration, 2003). Thus, one could argue that gender differences in behavioral self-handicapping and other self-defeating behaviors reflect a more general tendency among women to adhere to relevant societal norms.

There is a great deal of support for this possibility from a variety of sources. First, an examination of gender stereotype and sex role research reveals that women are expected to act in a communal manner, including being unselfish, dependent, and concerned with others. Men, on the other hand, are expected to be independent, assertive, and competent (Bem, 1974; Eagly & Steffen, 1984; Eagly & Wood, 1991). These differences are related to the fact that women are more likely to be in positions of less power, forcing them to be more dependent on others (Eagly & Wood, 1991). Future research should also examine whether status plays a more general role in effort beliefs and self-handicapping behavior. If it is lower social status that leads women to adhere to and internalize effort norms, then similar beliefs might be expected for other low status individuals. One would also predict that these individuals should show a similar disdain for certain types of self-handicapping behaviors.

Sex role differences are also reflected in personality descriptions. Women are more likely to describe themselves as agreeable and less likely to describe themselves as assertive than are men (Feingold, 1994). Taken together, these differences suggest that men may be given more latitude in choosing whether to adhere to social norms, particularly when doing so might conflict with being seen as assertive or competent, and/or that women are more willing to adhere to these norms. Similar conclusions can be drawn from research on social approval. Women are more likely to experience embarrassment (Miller, 1992), guilt (Tangney, 1990), higher fears of negative evaluation, and stronger motivations to avoid rejection (Miller, 1995) than are men.

Correlational data we have collected suggests that Worker scores are related to conscientiousness, agreeableness, and social desirability, and that these measures show corresponding gender differences (McCrea et al., 2006). Although it might be tempting to conclude based on these findings that these differences in valuing effort derive from a greater

pressure for women to adhere to norms and/or by a more general acceptance of norms by women, there is evidence from several sources that suggests this conclusion is not warranted. This same study (McCrea et al., 2006) indicated that gender differences in the Worker scale are not explained by differences in agreeableness, social desirability, or conscientiousness, strongly suggesting the present effects would not be explained by these constructs. It is also important to note that Study 2 and 3 illustrated that it is the personal value placed on effort, rather than more normative beliefs, that mediate the gender differences we observe. This finding suggests that the Worker scale is distinct from a more general pattern of adherence to social norms, and that differences in the internalization of the effort norm (as indexed by the Worker scale) are particularly crucial.

Moreover, Study 2 also illustrated that these beliefs about effort can be subservient to other concerns or norms, such as that of helping a friend. Thus, it is also not the case that individuals blindly adhere to these effort beliefs, independent of other important social norms. Indeed, this sensitivity to social context may explain why controlling for the Worker scale did not completely eliminate the gender effects in several of our studies, as there may be other relevant beliefs and norms invoked by the situation that come to bear in evaluating the attractiveness of a particular handicap. An important and exciting avenue for future research therefore concerns the interplay among these various normative and personal attitudes and the ubiquity of their application across a variety of different situations and social contexts.

## References

- Ablard, K. E., & Lipschultz, R. E. (1998). Self-regulated learning in high-achieving students: Relations to advanced reasoning, achievement goals, and gender. *Journal of Educational Psychology, 90*, 94-101.
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oakes, CA: Sage Publications.
- Arkin, R. M., & Oleson, K. C. (1998). Self-handicapping. In J. M. Darley & J. Cooper (Eds.), *Attribution and social interaction: The legacy of Edward E. Jones* (pp. 313-371). Washington, DC: American Psychological Association.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality & Social Psychology, 51*, 1173-1182.
- Baumeister, R. F., & Scher, S. J. (1988). Self-defeating behavior patterns among normal individuals: Review and analysis of common self-destructive tendencies. *Psychological Bulletin, 104*, 3-22.
- Baumgardner, A. H., & Levy, P. E. (1988). Role of self-esteem in perceptions of ability and effort: Illogic or insight? *Personality & Social Psychology Bulletin, 14*, 429-438.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting & Clinical Psychology, 42*, 155-162.
- Berglas, S., & Jones, E. E. (1978). Drug choice as a self-handicapping strategy in response to noncontingent success. *Journal of Personality & Social Psychology, 36*, 405-417.
- Cattell, R. B., & Cattell, A. K. S. (1961). *Test of "g": Culture fair intelligence test*. Champaign, IL: The Institute for Personality and Ability Testing.

- Cooper, H., Baumgardner, A. H., & Strathman, A. (1991). Do students with different characteristics take part in psychology experiments at different times of the semester? *Journal of Personality, 59*, 109-127.
- Courtenay, W. H., McCreary, D. R., & Merighi, J. R. (2002). Gender and ethnic differences in health beliefs and behaviors. *Journal of Health Psychology, 7*, 219-231.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354
- Dweck, C. S., Davidson, W., Nelson, S., & Enna, B. (1978). Sex differences in learned helplessness: II. The contingencies of evaluative feedback in the classroom, and III. An experimental analysis. *Developmental Psychology, 14*.
- Dweck, C. S., & Leggett, E. L. (1988). A social cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality & Social Psychology, 46*, 735-754.
- Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Journal of Personality & Social Psychology, 40*, 384-394.
- Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality & Social Psychology Bulletin, 17*, 306-315.
- Elliot, A. J., & Church, M. A. (2003). A motivational analysis of defensive pessimism and self-handicapping. *Journal of Personality, 71*, 369-396.
- Feingold, A. (1994). Gender differences in personality: A meta-analysis. *Psychological Bulletin, 116*, 429-456.

- Ferrari, J. R., & Tice, D. M. (2000). Procrastination as a self-handicap for men and women: A task-avoidance strategy in a laboratory setting. *Journal of Research in Personality, 34*, 73-83.
- Flamm, A. (2007). *When thinking means doing it: Prefactual thought in self-handicapping behavior*. Unpublished doctoral dissertation, University of Konstanz, Germany.
- Goodman, L. A. (1960). On the exact variance of products. *Journal of the American Statistical Association, 55*, 708-713.
- Harris, R. N., & Snyder, C. R. (1986). The role of uncertain self-esteem in self-handicapping. *Journal of Personality & Social Psychology, 51*, 451-458.
- Harris, R. N., Snyder, C. R., Higgins, R. L., & Schrag, J. L. (1986). Enhancing the prediction of self-handicapping. *Journal of Personality & Social Psychology, 51*, 1191-1199.
- Hendrix, K. S., & Hirt, E. R. (2006). *Stressed out over possible failure: The effects of induced self-regulatory focus on claimed self-handicapping*. Unpublished manuscript, Bloomington, IN.
- Hirt, E. R. (1993, May). *Behavioral self-handicapping in the social domain: Still a man's game?* Paper presented at the Midwestern Psychological Association, Chicago, IL.
- Hirt, E. R., Deppe, R. K., & Gordon, L. J. (1991). Self-reported versus behavioral self-handicapping: Empirical evidence for a theoretical distinction. *Journal of Personality & Social Psychology, 61*, 981-991.
- Hirt, E. R., McCrea, S. M., & Boris, H. I. (2003). "I know you self-handicapped last exam": Gender differences in reactions to self-handicapping. *Journal of Personality & Social Psychology, 84*, 177-193.

- Hirt, E. R., McCrea, S. M., & Kimble, C. E. (2000). Public self-focus and sex differences in behavioral self-handicapping: Does increasing self-threat still make it "just a man's game?" *Personality & Social Psychology Bulletin*, 26, 1131-1141.
- Jackson, D. N. (1974). Personality Research Form - Form E. In. Goshen, NY: Research Psychologists Press.
- John, O. P., & Srivastava, S. (1999). The Big-Five taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research (2nd ed.)* (pp. 102-138). New York: Guilford Press.
- Jones, E. E., & Berglas, S. (1978). Control of attributions about the self through self-handicapping strategies: The appeal of alcohol and the role of underachievement. *Personality & Social Psychology Bulletin*, 4, 200-206.
- Jones, E. E., & Rhodewalt, F. (1982). *The Self-Handicapping Scale*. Princeton, NJ: Princeton University.
- Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. In D. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 1, pp. 233-265). New York, NY: McGraw-Hill.
- Kimble, C. E., Funk, S. C., & DaPolito, K. L. (1990). The effects of self-esteem certainty on behavioral self-handicapping. *Journal of Social Behavior & Personality*, 5, 137-149.
- Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin*, 125, 470-500.
- Koch, K. A., Hirt, E. R., & McCrea, S. M. (2003). *Public self-focus and claimed self-handicapping*. Unpublished manuscript, Bloomington, IN.



- Kolditz, T. A., & Arkin, R. M. (1982). An impression management interpretation of the self-handicapping strategy. *Journal of Personality & Social Psychology, 43*, 492-502.
- Leary, M. R., & Shepperd, J. A. (1986). Behavioral self-handicaps versus self-reported handicaps: A conceptual note. *Journal of Personality & Social Psychology, 51*, 1265-1268.
- Luginbuhl, J., & Palmer, R. (1991). Impression management aspects of self-handicapping: Positive and negative effects. *Personality & Social Psychology Bulletin, 17*, 655-662.
- McCrea, S. M., & Flamm, A. (2006). [Anticipated regret, prefactual thinking and self-handicapping behavior]. Unpublished raw data.
- McCrea, S. M., & Flamm, A. (2007). *Planning to fail: Prefactual thought in self-handicapping*. Unpublished manuscript, Konstanz, Germany.
- McCrea, S. M., & Hirt, E. R. (2001). The role of ability judgments in self-handicapping. *Personality & Social Psychology Bulletin, 27*, 1378-1389.
- McCrea, S. M., Hirt, E. R., Milner, B. J., & Steele, N. L. (2006). *The Worker Scale: Gender differences in the personal value placed on effort*. Unpublished manuscript, Konstanz, Germany.
- Miller, R. S. (1992). The nature and severity of self-reported embarrassing circumstances. *Personality & Social Psychology Bulletin, 18*, 190-198.
- Miller, R. S. (1995). On the nature of embarrassability: Shyness, social evaluation, and social skill. *Journal of Personality, 63*, 315-339.
- Mirels, H. L., & Garrett, J. B. (1971). The Protestant work ethic as a personality variable. *Journal of Consulting & Clinical Psychology, 36*, 40-44.

- Mitchell, T. R. (1974). Expectancy models of job satisfaction, occupational preference and effort: A theoretical, methodological, and empirical appraisal. *Psychological Bulletin*, *81*, 1053-1077.
- Oleson, K. C., Poehlmann, K. M., Yost, J. H., Lynch, M. E., & Arkin, R. M. (2000). Subjective overachievement: Individual differences in self-doubt and concern with performance. *Journal of Personality*, *68*, 491-524.
- Poehlmann, K. M. (1994). *The phenomenology of overachievement*. Unpublished Unpublished Master's Thesis, The Ohio State University, Columbus, Ohio.
- Preacher, K. J., & Leonardelli, G. J. (2001, April 10). *Calculation for the Sobel test*. Retrieved June 20, 2003, from <http://quantrm2.psy.ohio-state.edu/kris/sobel/sobel.htm>
- Pyszczynski, T., & Greenberg, J. (1983). Determinants of reduction in intended effort as a strategy for coping with anticipated failure. *Journal of Research in Personality*, *17*, 412-422.
- Rhodewalt, F. (1990). Self-handicappers: Individual differences in the preference for anticipatory, self-protective acts. In R. L. Higgins, C. R. Snyder & S. Berglas (Eds.), *Self-handicapping: The paradox that isn't. The Plenum series in social/clinical psychology* (pp. 69-106). New York, NY: Plenum Press.
- Rhodewalt, F. (1994). Conceptions of ability, achievement goals, and individual differences in self-handicapping behavior: On the application of implicit theories. *Journal of Personality*, *62*, 67-85.
- Rhodewalt, F., & Davison, J. (1986). Self-handicapping and subsequent performance: Role of outcome valence and attributional certainty. *Basic & Applied Social Psychology*, *7*, 307-322.

- Rhodewalt, F., Saltzman, A. T., & Wittmer, J. (1984). Self-handicapping among competitive athletes: The role of practice in self-esteem protection. *Basic & Applied Social Psychology, 5*, 197-209.
- Rosenberg, M. (1965). *Society and adolescent self-image*. Princeton, NJ: Princeton University Press.
- Self, E. A. (1990). Situational influences on self-handicapping. In R. L. Higgins (Ed.), *Self-handicapping: The paradox that isn't. The Plenum series in social/clinical psychology* (pp. 37-68). New York, NY: Plenum Press.
- Sheppard, J. A., & Arkin, R. M. (1989). Determinants of self-handicapping: Task importance and the effects of preexisting handicaps on self-generated handicaps. *Personality & Social Psychology Bulletin, 15*, 101-112.
- Snyder, C. R., Ford, C. E., & Hunt, H. A. (1985). *Excuse-making: A look at sex differences*. Paper presented at the American Psychological Association, Los Angeles, CA.
- Stricker, L. J., Rock, D. A., & Burton, N. W. (1993). Sex differences in predictions of college grades from scholastic aptitude test scores. *Journal of Educational Psychology, 85*, 710-718.
- Strube, M. J., & Roemmele, L. A. (1985). Self-enhancement, self-assessment, and self-evaluative task choice. *Journal of Personality & Social Psychology, 49*, 981-993.
- Substance abuse and Mental Health Services Administration. (2003). *Results from the 2002 National Survey on Drug Use and Health: National Findings*. Rockville, MD.
- Swim, J. K., & Sanna, L. J. (1996). He's skilled, she's lucky: A meta-analysis of observers' attributions for women's successes and failures. *Personality & Social Psychology Bulletin, 22*, 507-519.

Tangney, J. P. (1990). Assessing individual differences in proneness to shame and guilt:

Development of the Self-Conscious Affect and Attribution Inventory. *Journal of Personality & Social Psychology*, 59, 102-111.

Tice, D. M. (1991). Esteem protection or enhancement? Self-handicapping motives and

attributions differ by trait self-esteem. *Journal of Personality & Social Psychology*, 60, 711-725.

Vroom, V. H. (1964). *Work and motivation*. New York: Wiley.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning:

Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82, 51-59.

## Appendix

### **Worker scale:**

1. I tend only to work as hard as I have to in my classes (reversed).
2. I try to devote my full effort to every class I take.
3. I pride myself in being a hard worker.
4. I push myself a lot to perform well academically.
5. I blow things off more than I should (reversed).
6. My grades are the result of effort and hard work.
7. I work hard to be successful at whatever I do.
8. I am proud to admit how hard I work to other people.

### **Academic Concern scale:**

1. I don't feel that I have to prove myself academically (reversed).
2. When I do poorly, I am afraid I am letting other people down.
3. I don't really worry much about what others think of me academically. (reversed)
4. I feel that I have to prove myself academically.
5. When I do poorly at a class, I feel that I have let myself down.
6. Others have high expectations from me academically.
7. I get really upset with myself when I don't perform well.

### **Prescriptive Effort Norm scale:**

1. Students who keep trying, even in the face of failure, should be admired.
2. I admire people who work hard.
3. Grades should be based in part on how much effort you put into a class.
4. People should strive to be the best at whatever they do.
5. Success due to effort is more meaningful than success due to ability alone.

### **Ability versus Effort Tradeoff scale:**

1. I would rather be envied for my ability than appreciated for my hard work.
2. I would be satisfied knowing I had the ability to do something but only lacked the desire to do it.
3. It is better to be seen as someone with potential than as a person who tried and failed.
4. In most classes, it is more important to appear intelligent than it is to learn the material.
5. I would rather be seen as competent than as a hard worker.
6. People respect ability more than they respect effort.

### Author Note

Sean M. McCrea, Department of Psychology, University of Konstanz; Edward R. Hirt and Bridgett J. Milner, Department of Psychological and Brain Sciences, Indiana University Bloomington. The authors wish to thank the many research assistants who helped collect and code the data for the present set of studies. Correspondence concerning this article should be addressed to Sean M. McCrea, Fachbereich Psychologie, Universität Konstanz, Universitätsstrasse 10, 78457 Konstanz, Germany. Electronic mail may be sent via the Internet to [Sean.McCrea@uni-konstanz.de](mailto:Sean.McCrea@uni-konstanz.de)

## Footnotes

1. Skeptical readers may debate whether Chris' behavior in our movie scenario can be unambiguously construed as self-handicapping. We took great pains to construct a scenario suggestive of the motive of self-handicapping (important performance relevant to self-esteem, high uncertainty, little if any prior preparation, clear evidence given that preparation would facilitate performance), and manipulated factors that should affect participants' construal of Chris' motives. However, we readily acknowledge that our scenarios (like any social behavior) were ambiguous and participants were indeed free to interpret Chris' behavior as they saw fit. The critical question in this research was whether actions that could serve as a behavioral self-handicap were evaluated positively or negatively by males and females. Thus, it was not crucial for us whether Chris "really" went to the movies in order to self-handicap or not, but rather how positively or negatively that behavior was perceived and what motives were ascribed to the target. The prediction was that females who self-handicapped in this way (i.e., lack of study effort) would be viewed more negatively than would males enacting the very same behavior.
2. Given that self-handicapping is expected to be highest for those low in Concern with performance and high in Self-doubt, we also examined whether there were gender differences on the product of these two scales. However, no such effect was observed ( $t < 1.29$ , ns).
3. We controlled for these variables to provide stronger evidence that lower reported effort was indicative of self-handicapping. However, gender differences on reported effort and evaluations of Chris, as well as the mediation of these effects by the Worker scale, remained significant when excluding these variables from the regression models.

4. It is important to clarify what metric we are using for behavioral self-handicapping here. Consistent with past work (Hirt et al., 1991; 2000), it is the *absence of a difference* in the amount of practice expended by individuals in the practice matters (handicap present) and practice doesn't matter (handicap absent) conditions that serves as evidence of behavioral self-handicapping. Participants expend very little practice effort in the practice doesn't matter condition, as it does not improve test performance; thus, there is a relative floor effect in the amount of practice observed in this condition. However, participants in the practice matters condition will practice significantly more unless they see reduced effort as an opportunity to self-handicap. Thus, in this paradigm, failure to expend effort in the practice matters condition denotes evidence of behavioral self-handicapping.
5. As in Study 1, we controlled for these variables to provide additional evidence that reduced practice effort was indicative of self-handicapping. When these variables were not included in the regression model, the Practice Instruction X Gender interaction remained significant, as did the mediation of the gender effect in the practice matters condition by the Worker scale.



Table 1

*Zero-order correlations between mediators (Study 1)*

	Academic Concern	Worker scale	Ability leads to success	Effort enjoy- ment	Effort leads to success	Concern with perform- ance	Self- doubt	Implicit theories	Self- esteem	SHS
Academic Concern	0.80									
Worker scale	.25†	0.86								
Ability leads to success	.07	-.20*	0.78							
Effort enjoyment	.02	.51‡	-.26†	0.92						
Effort leads to success	.14	.30‡	-.08	.55‡	0.92					
Concern with performance	.33‡	.34‡	-.10	.39‡	.29‡	0.80				
Self-doubt	.26†	-.17*	.20*	-.41‡	-.26†	-.14	0.83			
Implicit theories	-.02	.02	-.16	.09	.06	-.05	-.05	0.89		
Self-esteem	-.31‡	.28†	-.13	.33‡	.16	.14	-.64‡	.04	0.91	
SHS	.25†	-.54‡	.26†	-.51‡	-.20	-.15	.60‡	-.08	-.60	0.69

Note: \*  $p < .05$ , †  $p < .01$ , ‡  $p < .001$ . Reliabilities ( $\alpha$ ) are presented on the diagonal.

Table 2

*Gender effects on candidate mediators (Study 1)*

Measure	Men		Women		<i>t</i>	$\eta^2$
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>		
Academic Concern	24.95	5.62	27.01	5.03	2.30*	0.04
Worker scale	24.82	5.86	28.15	6.07	3.26†	0.07
Ability leads to success	17.37	5.23	17.66	3.94	<1	0.00
Effort enjoyment	52.84	11.31	53.04	10.68	<1	0.00
Effort leads to success	99.47	16.39	102.65	12.18	1.33	0.01
Concern with performance	40.79	8.16	40.75	6.61	<1	0.00
Self-doubt	25.5	7.30	29.3	6.64	3.28†	0.07
Implicit theories	13.30	3.79	13.33	3.78	<1	0.00
Self-esteem	56.82	10.73	50.18	11.03	3.60‡	0.08
SHS	60.16	12.87	60.24	10.84	<1	0.00

Note: \*  $p < .05$ , †  $p < .01$ , ‡  $p < .001$ .

Table 3.

*Mediational analysis for effects on reported study effort (Study 1)*

	Initial Model			Mediation Model		
	$\beta$			$\beta$	$t$	
Test importance	.236	2.65	<.01	.187	2.09	<.05
Fairness	-.189	1.67	.10	-.145	1.34	.18
Effort leads to success	.076	<1	ns	-.034	<1	ns
Test importance x Fairness	-.136	1.44	.15	-.111	1.24	.22
Test importance x Effort leads to success	.019	<1	ns	-.038	<1	ns
Fairness x Effort leads to success	-.047	<1	ns	.004	<1	ns
Test importance x Fairness x Effort leads to success	-.100	1.12	.26	-.065	<1	ns
Previous test score	.012	<1	ns	-.062	<1	ns
Satisfaction with previous score	-.021	<1	ns	-.015	<1	ns
Interest in course	.195	1.97	.05	.119	1.23	.22
Gender	-.196	2.23	<.05	-.019	<1	ns
Worker scale				.490	4.57	<.001
Academic Concern				-.162	1.68	.10
Self-esteem				-.168	1.34	.18
Self-doubt				-.088	<1	ns

$$R^2 = .21,$$

$$F_{\text{Change}}(11,117) = 2.78,$$

$$p < .01$$

$$R^2 = .34, \Delta R^2 = .13,$$

$$F_{\text{Change}}(4,110) = 5.36,$$

$$p < .001$$

Table 4.

*Mediational analysis for effects on evaluations of Chris (Study 1)*

	Initial Model			Mediation Model		
	$\beta$			$\beta$	$t$	
Test importance	-.252	2.84	<.01	-.225	2.50	<.05
Fairness	.112	<1	ns	.087	<1	ns
Effort leads to success	-.084	<1	ns	.023	<1	ns
Test importance x Fairness	.017	<1	ns	-.017	<1	ns
Test importance x Effort leads to success	.038	<1	ns	.040	<1	ns
Fairness x Effort leads to success	.122	1.38	.17	.106	1.21	.23
Test importance x Fairness x Effort leads to success	.001	<1	ns	-.054	<1	ns
Previous test score	-.071	<1	ns	-.001	<1	ns
Satisfaction with previous score	.001	<1	ns	.015	<1	ns
Interest in course	.068	<1	ns	.087	<1	ns
Gender	.233	2.68	<.01	.217	2.15	<.05
Worker scale				-.296	2.81	<.01
Academic Concern				.097	<1	ns
Self-esteem				.023	<1	ns
Self-doubt				.141	1.28	.20

$$R^2 = .21,$$

$$F_{\text{Change}}(11,119) = 2.83,$$

$$p < .01$$

$$R^2 = .30, \Delta R^2 = .10,$$

$$F_{\text{Change}}(4,112) = 3.90,$$

$$p < .01$$

Table 5

*Gender effects on candidate mediators within undesirable handicap condition (Study 2)*

Measure	Men		Women		<i>t</i>	$\eta^2$
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>		
Worker scale	21.89	6.90	25.18	5.75	3.28†	0.06
Prescriptive Effort Norm scale	20.23	3.70	21.64	2.11	3.02†	0.06
Ability vs. Effort Tradeoff scale	12.89	5.25	12.19	3.95	<1	0.01

Note: †  $p < .01$ .

Table 6.

*Mediational analyses for effects on target evaluation within undesirable handicap condition*

*(Study 2)*

	Initial Model			Mediation Model		
	$\beta$			$\beta$	$t$	
Gender	.258	3.33	<.01	.150	2.06	<.05
Worker scale				-.462	5.80	<.001
Prescriptive Effort Norm scale				.041	<1	ns

$R^2 = .07,$   
 $F_{\text{Change}}(1,156) = 11.11$   
 $p < .01$

$R^2 = .25, \Delta R^2 = .19,$   
 $F_{\text{Change}}(2,154) = 19.21,$   
 $p < .001$

Table 7

*Gender effects on candidate mediators within practice matters condition (Study 3)*

Measure	Men		Women		<i>t</i>	$\eta^2$
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>		
Worker scale	23.33	6.40	27.08	5.84	3.06†	0.09
Prescriptive Effort Norm scale	16.61	2.14	17.88	1.78	3.23†	0.10
Ability vs. Effort Tradeoff scale	13.37	4.75	12.51	4.46	1.13	0.01

Note: †  $p < .01$

Table 8

*Mediational analyses for effects on practice index within practice matters condition (Study 3)*

	Initial Model			Mediation Model		
	$\beta$			$\beta$	$t$	
Gender	-.370	3.94	<.001	-.249	2.76	<.01
Worker scale				.469	4.26	<.001
Prescriptive Effort Norm scale				-.051	<1	ns

$R^2 = .14,$   
 $F_{\text{Change}}(1,98) = 15.55,$   
 $p < .01$

$R^2 = .31, \Delta R^2 = .18,$   
 $F_{\text{Change}}(2,96) = 12.43,$   
 $p < .001$

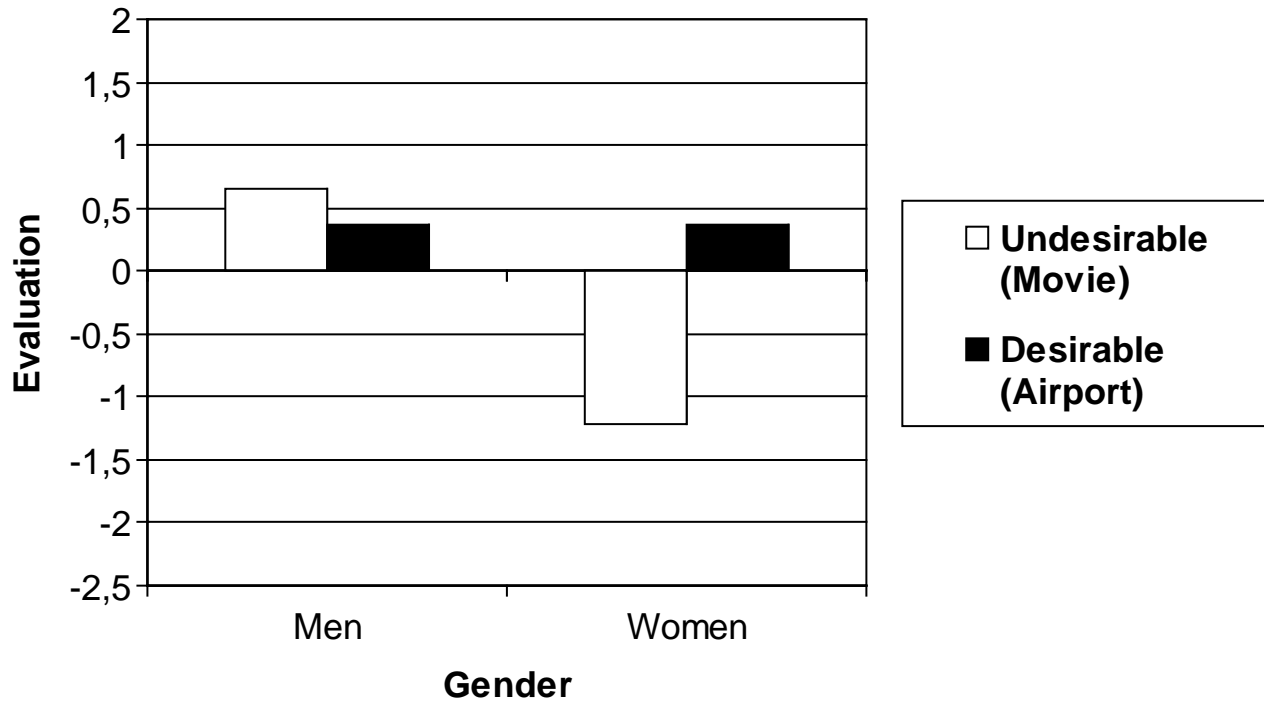


*Figure 1.* Gender by Type of behavior interaction (Study 2).

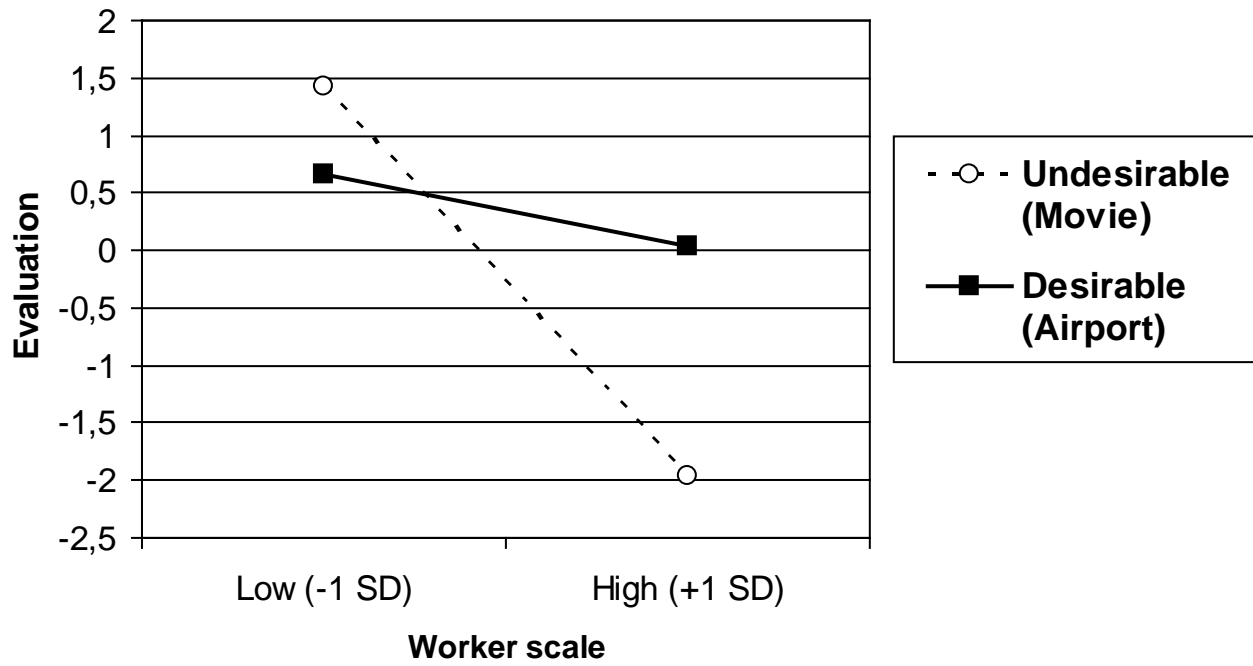
*Figure 2.* Worker scale by Type of behavior interaction (Study 2).

*Figure 3.* Adjusted practice index scores (Study 3).

### Gender x Type of behavior



### Worker scale x Type of behavior



### Practice Effort

