PERGAMON

Personality and Individual Differences 33 (2002) 587–605

PERSONALITY AND INDIVIDUAL DIFFERENCES

www.elsevier.com/locate/paid

What we really know about our abilities and our knowledge

Phillip L. Ackerman*, Margaret E. Beier, Kristy R. Bowen

School of Psychology, Georgia Institute of Technology, Room 130 Psychology Building, 274 5th Street, MC 0170, Atlanta, GA 30332-0170, USA

Received 6 April 2001; received in revised form 27 August 2001

Abstract

Recently, it has become popular to state that "people hold overly favorable views of their abilities in many social and intellectual domains" [Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognising one's own incompetence lead to inflated self-assessments. *Journal of Personality & Social Psychology*, *77*(6), (1999) 1121]. Research that supports this point tells only half of the story—in a manner documented by Cronbach's [Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, (1957) 671] classic article on the "two disciplines of scientific psychology." That is, the recent research has only documented the experimental side of the scientific divide (which focuses on means and ignores individual differences). The current paper shows that research from the other side of the scientific divide, namely the correlational approach (which focuses on individual differences), provides a very different perspective for people's views of their own intellectual abilities and knowledge. Previous research is reviewed, and an empirical study of 228 adults between 21 and 62 years of age is described where self-report assessments of abilities and knowledge are compared with objective measures. Correlations of self-rating and objective-score pairings show both substantial convergent and discriminant validity, indicating that individuals have both generally accurate and differentiated views of their relative standing on abilities and knowledge. © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Abilities; Knowledge; Self-rating of ability; Self-concept

1. Introduction

A fairly unflattering picture of people's skill to accurately judge their own abilities has recently emerged from the psychology literature. The general tendency for people to be inaccurate—either unaware of their own incompetence, or likely to underestimate their skill—has been highlighted in the experimental literature (Fischhoff, Slovic, & Lichtenstein, 1977; Kruger, 1999; Kruger &

* Corresponding author. Fax: +1-404-252-3061/+1-404-894-6904.
E-mail address: phillip.ackerman@psych.gatech.edu (P.L. Ackerman).

PII: S0191-8869(01)00174-X

Dunning, 1999). For example, Kruger and Dunning (1999) hypothesized that incompetent individuals "will tend to grossly overestimate their skills and abilities" (p. 1122). According to this body of research, various factors may moderate or contribute to the tendency to misjudge one's abilities. These include knowledge of the domain being tested (Lichtenstein & Fischhoff, 1977; Shaughnessy, 1979), the difficulty of the domain (Kruger, 1999), the specificity/ambiguity of the ability being evaluated (Dunning, Meyerowitz, & Holzberg, 1989), knowledge of and specificity of the referent group (Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995), desirability of the trait being evaluated (Alicke, 1985), gender differences (Lundeberg, Fox, Brown, & Elbedour, 2000; Lundeberg, Fox, & Puncochar, 1994), cultural differences (Lundeberg et al., 2000; Yates, Lee, & Bush, 1997) and ability (Maki, Jonas, & Kallod, 1994; Moreland, Miller, & Laucka, 1981). Irrespective of these different factors, the common theme expressed is that the ability to accurately judge one's own ability is sadly lacking.

Kruger and Dunning (1999) recently studied people's tendencies to misjudge their abilities across the domains of humor, logical reasoning and English grammar. In these studies participants were asked to provide an estimated percentile rank of their ability compared to their peers (peers were defined as average undergraduates or, in some cases, average "other" students in their classes) before a test of domain knowledge. After being tested, participants were asked to estimate their test performance compared to their peers by providing a percentile rank for their test performance. Participants were also asked to report the number of items they felt they answered correctly out of the total number of items on the test. Across all domains, these researchers found that those in the bottom quartile for performance were most likely to overestimate their ability. These participants seemed to inflate their self-reported percentile rank before they took the test and overestimated their perceived test performance after they took the test. Kruger and Dunning (1999) called this the "dual burden" of those with low ability. "Not only do they reach erroneous conclusions and make unfortunate choices, but their incompetence robs them of the ability to realize it" (p. 1121). In their study, top performers also seemed likely to misjudge their abilities. In contrast to low performers, top performers appeared to underestimate their abilities compared with their peers.

Researchers have remarked that the true relationship between objective ability and an individual's perception of his/her own ability may be obscured by the data-analytic techniques that are used to assess it. Indeed, the results found by Kruger and Dunning (1999) might be accounted for by a statistical artifact, regression to the mean. To illustrate this point, we examined information provided in Study 3, Phase 1 of Kruger and Dunning (1999). This study compared the relationship between self-reported percentile rank (compared to others in the class) on a test of English grammar and objective test scores (also in percentile rank). Kruger and Dunning (1999) found a significant but small positive correlation between perceived test score and actual test score, $r=0.19$. Further, Kruger and Dunning (1999) found that when participants were grouped into quartiles based on their performance on the objective test, those in the bottom quartile were likely to significantly overestimate their ability while those in the top quartile were likely to significantly underestimate their ability.

To understand whether this effect could be accounted for by regression to the mean, we simulated this analysis using two random variables (one representing objective knowledge and the other self-reported knowledge) and 500 observations (representing an N of 500). As in the Kruger and Dunning (1999) comparison, these random variables were correlated $r=0.19$. The observations were then divided into quartiles based on the simulated scores for the objective knowledge variable ($n=125$ observations per quartile). Simulated self-report and objective knowledge were

then compared by quartile. As can be seen in Fig. 1, the plotting of simulated data for 500 subjects resulted in exactly the same phenomenon reported by Kruger and Dunning (1999)—an overestimation for those in the lowest quartile and an underestimation for those in the top quartile. Further analysis comparing the means of self-report and objective knowledge for each quartile revealed that the difference between the simulated self-reported ($M=-0.21$) and objective ($M=-1.22$) scores for the bottom quartile was significant $t\ (124)=-10.09$, $P<0.001$ (which would be "interpreted" as overestimation of performance). The difference between simulated self-reported ($M=0.27$) and objective ($M=1.36$) scores for the top quartile was also significant $t(124)=11.09$, $P<0.001$, ("interpreted" as underestimation by top performers). This illustration demonstrates the measurement problems associated with interpreting statistical significance when two variables are compared across groups selected for performance on one of the variables, and there is a low correlation between the two variables.

Kruger and Dunning (1999) discussed two reasons why they believe the effects found in their study reflect a lack of metacognitive ability and not a statistical artifact. First, they stated that the magnitude of miscalibration would be the same for both top and bottom quartiles if the effect were due to regression to the mean—and they found that the bottom quartile tended to overestimate their performance more than the top quartile underestimated their performance. However, across experiments, participants generally perceived their rank around the 66th percentile, on average.

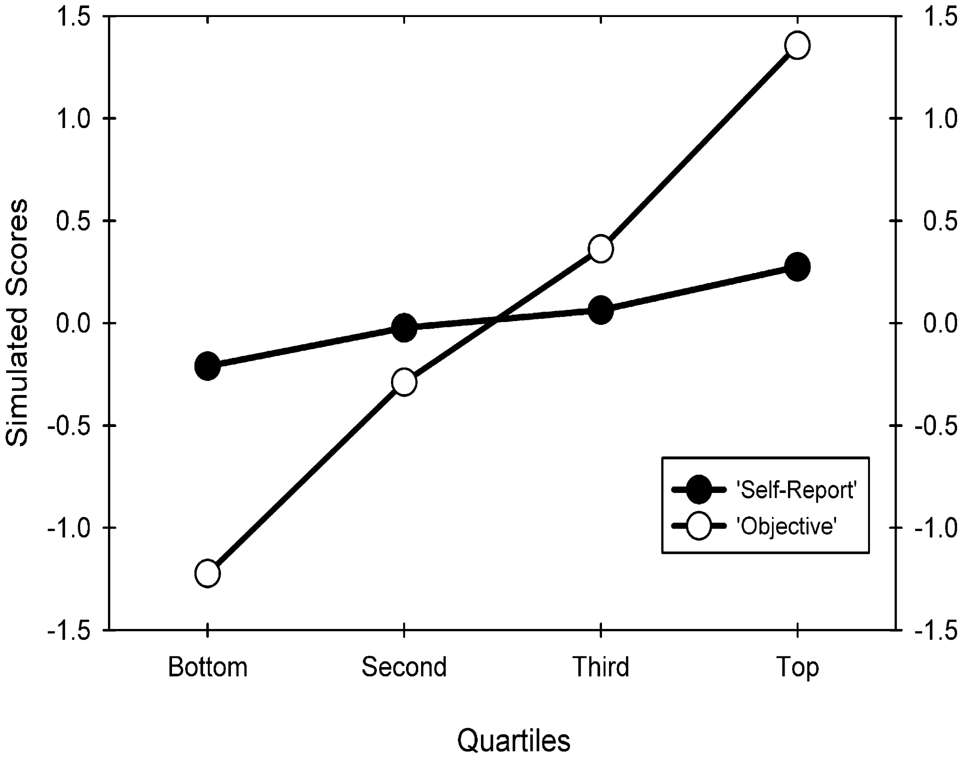

Fig. 1. Results representing simulated self-report scores as a function of simulated objective performance (500 observations). Quartiles established from simulated 'objective' performance scores. Mean scores for 'objective' and 'self-report' data are plotted. Each quartile represents 125 observations.

Dividing the sample into quartiles based on their performance (regardless of the range of raw scores) would lead to a greater miscalibration for the bottom quartile than the top (i.e. since the bottom quartile will necessarily be further from the 66th percentile than the top). Second, Kruger and Dunning (1999) stated that their findings were not solely the result of regression to the mean because they were able to improve the metacognitive skill of the participants through training. However, overestimation and underestimation still occurred for these groups even after training—regression to the mean may still be an issue in this analysis.

Even with the possibility of statistical artifacts like regression to the mean, the experimental literature describes both skilled and unskilled performers as likely to misjudge their abilities. The research we present here provides a broad and more optimistic interpretation of self-assessment of ability. The current study employs measures of self-concept, self-estimates of ability and self-estimates of knowledge—each measure is discussed below.

1.1. Self-concept

There is a substantial history to the construct of self-concept (e.g. James, 1890/1950). Self-concept is a relatively broad construct, and some researchers have partitioned self-concept across academic, social, emotional and physical domains. Marsh (1990), segmented self-concept into four categories: Academic, Social, Emotional, and Physical self-concepts. These conceptualizations use judgments of competencies in specific areas as their basis. Item scores are then aggregated for a broader conceptualization of self-concept. Gender differences in self-concept have also been reported (Marsh, 1990; Marsh & Yeung, 1998) and are generally consistent with gender role and stereotypes about interest and ability (i.e. women typically aligned with verbal abilities and humanities, men typically aligned with math, science and technology). Relatively little research has been reported on the congruence between self-concept and objective measures, though there has been research on the congruence between self-concept and others' ratings (e.g. Shavelson, Hubner, & Stanton, 1976) that have reported self—other agreement for specific self concepts, but less agreement as the traits or abilities one is asked to rate increase in breadth.

Although it is certainly true that some individuals have distorted self-concepts, it would be *erroneous to believe that people are generally overconfident or generally underconfident of their abilities in this sense*. In the physical self-concept domain, consider the following statements: (1) I could run a mile in under 4 min; (2) I could lift a 400 lb weight; (3) I could successfully play in a world-class football match. Although a small portion of the population is capable of each of these actions, few other individuals would be likely to be sufficiently overconfident of their abilities to positively endorse these items. Similarly, in the occupational or academic domains, if one were to ask about self-concept for "brain surgery" or "nuclear physics"—most individuals would correctly indicate that they do not have the ability to carry-out such activities successfully. Conversely, if we ask "I could drive to the grocery store" or "I could brush my teeth," it would be surprising if a significant portion of a participant sample would not indicate high agreement with these items. Although the above appear to be extreme examples, and therefore hold no ambiguity for the respondent, the point is that the likelihood of *overestimating* or *underestimating* one's own abilities depends much more on the ambiguity of the statement: (1) where the activity or aptitude is unknown to the individual—and so must be inferred from other information (e.g. "I could learn to juggle"); or (2) the domain to be considered is over-broad (e.g. "intelligence"). This suggests

that improvements in self-report/objective calibration can be obtained if the measures are sufficiently *specific*, rather than general or ambiguous. In the personality domain, the importance of specificity for self-assessment of traits has been examined by Dunning et al. (1989). For example, people were more likely to show a self-serving assessment of the trait when it was ambiguous (e.g. how sensitive are you) than they were when the trait was more specific (e.g. how punctual are you). In the abilities domain, it may be that the more appropriate consideration is a proper match between the generality/specificity of the self-report measure and the generality/specificity of the criterion objective measure—that is, what Wittmann and Süß (1999) referred to as "Brunswik symmetry."

In a study of the determinants of self-efficacy for learning and performance, Kanfer, Ackerman, and Heggestad (1996) found that a self-report measure of "academic self-concept" had a substantial pattern of correlations with Spatial Ability ($r = 0.439$) and Math Ability ($r = 0.459$). Moreover, math/science self concept provided a direct link to initial task self-efficacy, whereas ability and need for Achievement provided only *indirect* linkages—through their influence on self-concept. That is, self-concept appears to be a more important determinant of task-specific self-efficacy than do objective measures of ability.

1.2. Self-estimates of ability

Although self-concept measures occasionally ask the respondent to rate himself/herself on a particular ability (e.g. Math Ability), we believe that it is possible to differentiate between self-report endorsements of statements of particular skills (on an absolute scale) and self-report statements of broad abilities—especially in a normative context (e.g. I am above average in Math Ability). Kruger (1999) suggested that individuals underestimate their abilities compared to others when the task is hard, and overestimate their abilities when the task is easy. He attributes this phenomenon to the tendency of inadequately taking into consideration the skills of the referent group when making the comparison. An example of this phenomenon given by Kruger (1999) is the tendency for people to underestimate their ability to play chess compared to others and overestimate their ability to drive a car. In this normative context, the historical evidence on the congruence between self-estimates of ability and objectively determined ability is somewhat mixed (e.g. DeNisi & Shaw, 1977; Goff, 1994; Kelso, Holland, & Gottfredson, 1977; Levine, Flory, & Ash, 1977), but there is substantial evidence that self-estimates of ability provide convergent and discriminant relations with objective ability measures (e.g. see Ackerman, 1997 for a review). For convergent validity evidence, Ackerman, Kanfer, and Goff (1995) found that self-estimates of math/spatial ability and verbal ability correlated $r = 0.58$ and $r = 0.42$ with objective test composites of the math/spatial and verbal abilities, respectively. For discriminant validity, self-estimates of math/spatial ability correlated $r = -0.07$ with objective verbal ability, and self-estimates of verbal ability correlated $r = -0.10$ with objective math ability.

As stated by Ackerman (1997) "the correlations between self-ratings of ability and objective ability measures are not as large as the reliabilities of the separate measures [and thus, discrepancy (difference) scores might have sufficient reliability for further investigation]; and to date, there has been little interest in the literature in pursuing the meaning of objective/self-report discrepancies, especially as they relate to personality and other traits (Cantor & Norem, 1989). Such an integration between objective ability measures, self-estimates of ability, and personality characteristics, may yield some potentially important insights into the mechanisms of motivation and

choice behaviors (e.g. it is quite possible that an individual's choice of whether to engage a learning task is more substantially predicated on self-estimates of ability than objectively determined ability, or that self-objective discrepancy values are related to personality and self-esteem traits.)" (Ackerman, 1997, pp. 181–182).

1.3. Self-estimates of knowledge

There is a literature on the calibration of an individual's 'feeling of knowing' and the individual's actual performance on particular tasks (e.g. see Miner & Reder, 1994 for a review). However, that particular field of inquiry has not yielded sufficient information on the issues of individual differences in calibration. From the individual-differences perspective, one study (Rolfhus & Ackerman, 1996) examined the structure of self-reported knowledge across 32 different domains, and the correspondence with objective measures of abilities. These authors found that self-estimates of knowledge in Math and Physical Sciences, and in Technology had substantial correlations with objective measures of math, spatial, and mechanical abilities. Self-reported knowledge in Humanities and Social Sciences failed to correlate substantially with measures of math, spatial, mechanical, or verbal abilities. However, there was no direct comparison in that study between self-reported knowledge and objective knowledge measures.

2. The current study

The current study highlights the need to evaluate the individual-differences basis for correspondence between self-evaluations and objective measures of ability and knowledge. In particular, self-report measures of self-concept, self-estimates of ability, and self-report knowledge were obtained, and are compared against objective measures of ability and knowledge. The data for the current study are from a larger study of individual differences in ability and personality relations to knowledge across the adult lifespan (Ackerman, 2000). Although analyses of the objective ability and knowledge data have been reported elsewhere, the current study addresses the self-report data that have not been reported in the literature. Because this study contains both self-reported and objective ability and knowledge data, it is possible to analyze a variety of issues related to the congruence between people's views of their abilities and their objectively determined abilities. In addition, we will address both gender differences and age effects on these measures.

[Note: The method section has substantial overlap with Ackerman (2000). As such, the objective ability and knowledge measures are only summarized here. The self-report measures of ability, self-concept and knowledge measures are fully described, as they do not appear in the other article.]

3. Method

3.1. Participants

Participants were recruited through flyers and newspaper advertisements at (a large state university). Two-hundred and thirty individuals participated in the study; the data from two participants were

excluded because of a failure to follow instructions. Of the 228 participants in the final sample, 78 were men. The sample had a mean age of 34.2 years, and S.D. age of 10.6 years (105 participants between 20 and 29 years old; 52 between 30 and 39; 46 between 40 and 49; 20 between 50 and 59; and 5 between 60 and 62). Participants were required to have attained at least a BA/BS level of education, have normal or corrected-to-normal hearing, vision, and motor coordination. Highest educational attainment reported by the participants was BA/BS only (75%), MA/MS (19%), JD (2%), MD or DVM (1%), PhD (4%).

3.2. Apparatus

Paper-and-pencil ability tests were administered in a laboratory room with up to 16 participants at a time. Knowledge tests were administered using IBM-compatible computers. The participants were instructed to complete the take-home questionnaire at home in a quiet, undisturbed environment.

3.3. Measures

3.3.1. Ability battery

The ability battery contained 13 tests selected to provide assessment of fluid intellectual ability (Gf) and crystallized intellectual ability (Gc). To assess Gf, seven tests were administered. First included was the Raven Advanced Progressive Matrices test. Other reasoning tests were also included: Spatial Analogy, Number Series, and Diagraming Relations, along with an auditory Number Span test and two math problem solving tests, called Problem Solving and Necessary Facts.

To assess Gc, several verbal content measures were included: a test of Vocabulary, the Nelson-Denny Reading Speed score, the Nelson-Denny Comprehension Test and the Comprehension test from the Multidimensional Aptitude Battery (MAB), a fluency test called Word Beginnings, the [Similarities] test from the MAB, and a written adaptation of the Wechsler Adult Intelligence Scales-Revised (WAIS-R) Information Scale. (For additional details, see Ackerman, 2000).

3.4. Knowledge scales

Development of the knowledge test battery is described in Rolfhus and Ackerman (1999). The battery is based on a series of tests provided by the College Board [Advanced Placement (AP) and College Level Experience Program (CLEP) tests]. Extensive item-development, testing and validation was undertaken with over 700 individuals. The result is a series of objective (multiple choice) knowledge tests, administered in a power format. Items covered the entire range of difficulty. The easiest items were designed to be easily completed by all participants; the most difficult items were intended to be beyond reach for all participants (except perhaps by career experts in that domain). Tests began with the easiest item in a domain. Items were presented in order of difficulty. When an individual responded incorrectly to three consecutive items, he/she was moved to the next domain. An individual's score was the number of items completed correctly.

Eighteen domains were selected from a set of twenty used in Ackerman and Rolfhus (1999) and Rolfhus and Ackerman (1999). The tests represent four broad, but somewhat overlapping domains: (1) *Science* (Astronomy, Biology, Chemistry, Electronics, Physics, Psychology, and Technology), (2) *Civics* (US Government, US History, Economics, Geography, and Western

Civilization), (3) *Humanities* (US Literature, Art, Music, and World Literature); and (4) *Business* (Business/Management and Law).

3.5. Self-report scales

The self-report measures including demographic background, knowledge, self-concept, self-estimates of abilities, and other measures of personality and interests not reported here were contained in the questionnaire. Table 1 provides descriptive information for the self-report scales.

3.5.1. Self-estimates of abilities scales

For self-ratings of ability, see Ackerman et al. (1995), Ackerman and Rolfhus (1999), Kanfer et al. (1996). The questionnaire items represented five scales of broadly described abilities, aptitudes, or knowledge, in the domains of verbal, math, mechanical, clerical, and self-management (e.g. "Math Ability", "Reading Ability", "Scientific Ability"). The instructions directed the participant to "Rate yourself on each of the following traits as you really think you are when compared with other persons your own age. Give the most accurate estimate of how you see yourself. Avoid rating yourself the same on each ability or trait." Response options were presented in a Likert-type scale. Score was the average response made to the questions for each scale.

3.5.2. Self-concept scales

Self-concepts for competencies and aptitudes were assessed with an adaptation of scales initially developed by Goff (1994), and later validated by and expanded by Ackerman et al. (1995), Kanfer et al. (1996), see also Ackerman and Rolfhus (1999). The instructions direct participants to "Consider whether you have the skill or ability, keeping in mind that most people vary in the kinds of skills and abilities that they have." Each item provides a specific reference to a particular skill or ability. A few examples: "I can recognize correct English usage (that is, grammar and punctuation)." "I understand science articles or programs." "I understand and can use things presented as figures, such as maps or blueprints." Response options were presented in a Likert-type scale. Four self-concept scales were administered (verbal, math, spatial, and science). Score was the average response made to the questions.

3.5.3. Self-report knowledge scales

There was a parallel set of 18 self-report knowledge scales for the 18 objective knowledge scales. Each self-report scale was composed of six items with a specific reference to the kinds of knowledge required on the respective objective knowledge scale. The instructions direct participants to: " ...rate your knowledge about each topic. That is, if you know very little about a topic, you would write in a 1 or 2 for the item. If, however, you know a great deal about a topic, you would write in a 5 or a 6 for the item." For example, the US Government scale included the following questions: "(1) The separation of powers between the representative, judicial and executive branches of government. (2) The role of special interest groups in the legislative process. (3) Duties of various government officials, such as the Secretary of State and the Attorney General. (4) The historic separation of church and state in the US. (5) The electoral process. (6) Term limits and lengths for elected officials." Response options were presented in a Likert-type scale. Score was the average response made to the questions.

3.6. Procedure

"In the first session (3 h with breaks), participants completed an aptitude battery of 13 tests designed to assess Gf and Gc, all using paper-and-pencil procedures. The participants were given

Table 1
Descriptive statistics for self-report measures

Scale	Total no. of items	Mean	S.D.	α
Self-estimates of ability				
Verbal	4	5.41	0.89	0.72
Math	3	4.40	1.28	0.81
Clerical	2	4.84	1.28	0.88
Mechanical	2	4.33	1.12	0.80
Self-concept				
Verbal	6	4.09	0.53	0.74
Math	6	3.33	0.94	0.92
Spatial	6	3.63	0.75	0.85
Science	6	3.50	0.85	0.92
Knowledge				
Science				
Astronomy	6	2.80	1.48	0.93
Biology	6	3.76	1.70	0.94
Chemistry	6	2.98	2.16	0.97
Electronics	6	1.29	1.36	0.92
Physics	6	2.11	1.74	0.95
Psychology	6	2.78	1.55	0.93
Technology	6	2.21	1.47	0.88
Civics				
Economics	6	3.20	1.51	0.92
Geography	6	4.26	1.25	0.89
Western Civilization	6	3.16	1.40	0.92
US Government	6	3.75	1.35	0.93
US History	6	3.67	1.40	0.91
Humanities				
Art	6	2.99	1.60	0.94
Music	6	2.99	1.45	0.90
US Literature	6	2.70	1.36	0.90
World Literature	6	2.28	1.24	0.86
Business				
Business/Management	6	3.44	1.60	0.94
Law	6	3.94	1.39	0.92

Response Scales: Knowledge (0 = Not at all knowledgeable to 7 = Extremely knowledgeable); Self-Concept (1 = Strongly disagree to 6 = Strongly agree); Self-Estimates of Ability (1 = Extremely low to 7 = Extremely high); $N = 228$. α = Cronbach's alpha measure of internal consistency reliability.

the questionnaire after the first session, and instructed to complete it and return it at the beginning of the second session. The second session followed the first by at least 24 h. The second session was composed entirely of the computerized knowledge tests. The knowledge tests were given in a prespecified order, with the starting point in the test sequence assigned at random to each participant. The test presentation order was arranged so that closely related domains were spread out (e.g. the physical sciences domains were not all administered sequentially). Participants received $100 for participation at the conclusion of the second session" (Ackerman, 2000).

4. Results

We describe the results in four sections: (1) Descriptive statistics for the self-report measures; (2) Correlations between self-estimates of knowledge, self-concept and objective ability; (3) Relationships between self-reported knowledge and objective knowledge; and (4) Relationships of objective and self-report knowledge by gender, age and major field of study.

4.1. Descriptive statistics on self-report measures

Table 1 shows the means, standard deviations and internal consistency reliabilities for the self-report measures of knowledge, self-concept, and self-estimates of abilities. On first glance, the self-estimates of abilities (a normative scale) are most closely illustrative of the so-called "Lake Wobegon Effect"—that is the general sense of most people to think that they are "above average." The mean scores on these measures ranged from $M=4.40$ to 5.41, which indicates that the average response to Math Ability was above average 4.40 (4 = "Average"), and the highest average scores were in Verbal Ability 5.41 (5 = "Moderately Above Average" and 6 = "Very High"). It is useful to note, though, that all of these participants had achieved the BA/BS level of education, and that 25% of the participants had achieved a postgraduate degree (either masters, professional, or PhD). So, it is not clear whether it is entirely fair to consider this an example of the Lake Wobegon effect, since these participants were indeed arguably "above average" in intellectual abilities.

When the questions are more specific (as in the Self-Concept and Self-Report of Knowledge) rather than general (as in the Self-Estimates of Ability), the responses of the participants are much more conservative. For self-concept, the questions pertained to specific aptitudes and skills (e.g. "I have always done well in math courses"). On these scales, the average scores ranged from $M=3.33$ to 4.09, where 3 = "Slightly Disagree" and 4 = "Slightly Agree," and 3.5 would represent a neutral statement (neither disagree nor agree). With this as a reference point, the Verbal Self-Concept showed a mean value significantly above the theoretical neutral point of 3.5, [$M=4.09$, $t(227)=16.86$, $P<0.01$]. A smaller, but still significant effect was found for the measure of Spatial Self Concept [$M=3.63$, $t(227)=2.71$, $P<0.01$]. In contrast, the mean value for the Math Self Concept measure was significantly below the neutral point of 3.5 [$M=3.33$, $t(227)=-2.81$, $P<0.01$], and the measure of Science Self Concept was not significantly different from the neutral point of 3.5.

For the self-report knowledge scales, the response format ranged from 0 = "Not at All Knowledgeable" to 7 = "Extremely Knowledgeable." There was a wide range of average scores on these

scales, from Electronics ($M=1.29$) to Geography ($M=4.26$) and substantial variability (the largest interindividual variance was found for Chemistry, S.D. = 2.16), where 1 = "Very Slightly Knowledgeable" and 4 = "Moderately Knowledgeable"). In every domain except for Geography, there were participants who rated themselves as essentially "Not at all Knowledgeable" about the field. In addition, these 18 scales showed "positive manifold," that is, ubiquitous positive intercorrelations. Only one correlation among the 153 unique correlations was significantly negative (between Business and Chemistry knowledge, $r=-0.148$, $P<0.05$), indicating that there is a presence of some underlying general factor of self-reported knowledge.

4.2. Self-concept, self-estimates of ability, and objective measures of ability and knowledge

Table 2 presents the intercorrelations among measures of self-concept, self-estimates of ability on the one hand and objective ability and knowledge measures on the other hand. The intercorrelations among the self-report measures indicate both convergent and discriminant validity, concordant with prior literature (Ackerman et al., 1995). That is, self-concept measures and corresponding self-estimates of ability (i.e. Verbal Self Concept vs. Self-Estimate of Verbal Ability and Math Self-Concept vs. Self-Estimate of Math Ability) show high concordance, even though

Table 2
Self-concept, self-estimates of ability and objective ability and knowledge composite intercorrelations[a]

	Verbal self-concept	Verbal self-estimate	Math self-concept	Math self-estimate	Spatial self-concept	Mechanical self-estimate	Clerical self-estimate
Self-report							
Verbal self-concept							
Verbal self-estimate	*0.674***						
Math self-concept	−0.129	−0.216**					
Math self-estimate	−0.157*	−0.154*	*0.858***				
Spatial self-concept	−0.075	−0.148*	0.385**	0.329**			
Self-estimate of mechanical ability	−0.194**	−0.088	0.115	0.157*	*0.464***		
Self-estimate of clerical ability	0.278**	0.283**	−0.075	−0.061	−0.058	−0.160*	
Objective							
Fluid intelligence (Gf)[b]	0.006	−0.053	*0.505***	*0.468***	*0.256***	0.051	−0.091
Crystallized intelligence (Gc)[b]	*0.279***	*0.355***	0.188**	0.176**	0.061	−0.085	−0.119
Science knowledge	−0.090	0.004	*0.441***	*0.410***	*0.261***	0.154*	−0.242**
Civics knowledge	0.097	*0.272***	0.176**	0.185**	0.102	−0.018	−0.167*
Humanities knowledge	*0.290***	*0.446***	−0.091	−0.062	−0.061	−0.124	0.030
Business knowledge	0.149*	0.280**	0.094	0.111	−0.052	−0.064	−0.005

Convergent validity indices shown in italics.
[a] Self-estimate = self-estimate of ability.
[b] Correlation between fluid intelligence (Gf) and crystallized intelligence (Gc) ($r=0.589$**).
* $P<0.05$.
** $P<0.01$.

the measures differ in both content (self-concept questions refer to specific knowledge and skills and self-estimates of ability refer to broad abilities). The verbal measures correlated $r=0.674$ with one another, and the math measures correlated $r=0.858$ with one another. Again, similar to previous research (Marsh & Yeung, 1998), the correlations between verbal measures and math measures are small and negative, indicating that the participants appear to slightly polarize their self-reported abilities, *even though verbal and math abilities are positively correlated with one another in the population* (i.e. if the participant thinks that he/she is high on verbal abilities, then he/she tends to discount his/her standing on math abilities, and vice versa). Although not completely overlapping, there was a substantial correlation between Spatial Ability self-concept and Self-Estimate of Mechanical Ability ($r=0.464$).

With respect to correspondence between self-reported and objective measures, positive and significant correlations were found between Gf and math and spatial self-reported abilities, and similar, but smaller correlations were found between Gc and verbal self-reported abilities. Non-significant correlations were found between self-reported verbal abilities and Gf, and small ($r<0.20$) but significant correlations were found between the math and spatial self-reported abilities and Gc, *even though Gf and Gc correlated* $r=0.589$. For the knowledge measures, composite scores were created for the four domains (Science, Civics, Humanities, and Business). Self-reported Math, Spatial, and Mechanical abilities correlated positively and significantly with

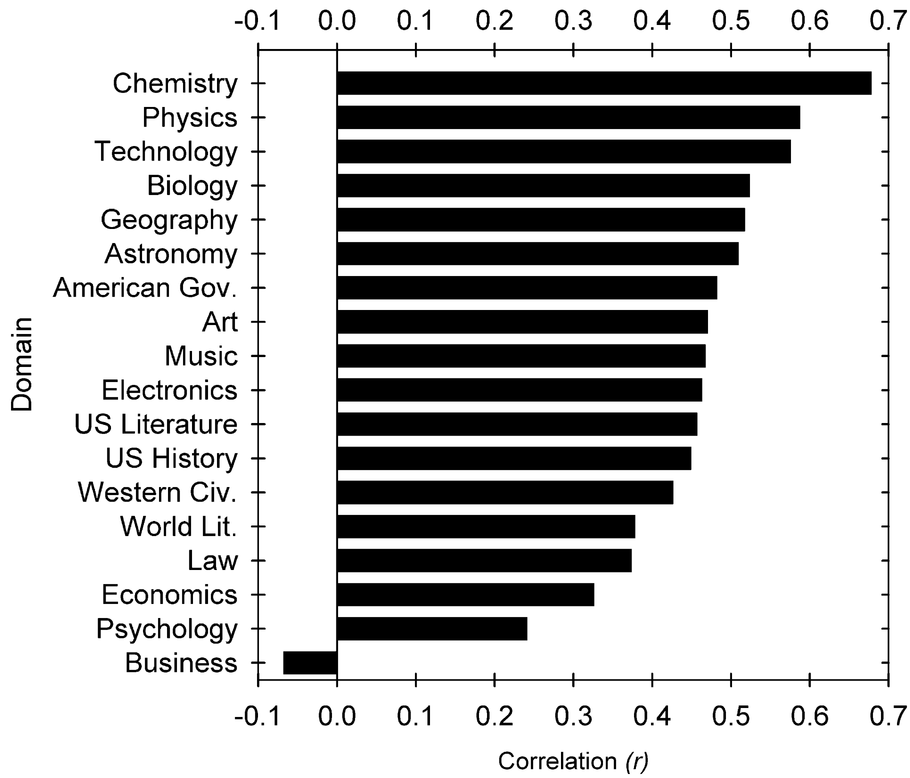

Fig. 2. Correlations between self-report and objective knowledge measures, sorted by magnitude of correlation. Gov., Government; Lit, Literature.

Science knowledge, while Humanities, Civics, and Business showed generally positive and significant correlations with self-reported verbal abilities. Only self-reported Clerical ability showed significantly negative correlations with objective knowledge, for both Science and Civics.

4.3. Self-estimates of abilities and knowledge vs. objective abilities and knowledge

It is first important to note that the self-report knowledge measures were administered before any of the objective knowledge measures. Fig. 2 shows the raw correlations between the self-ratings of knowledge and objective knowledge measures, ordered by the magnitude of the correlations. These results provide an illuminating perspective on the correspondence between individuals' perceived degree of knowledge and their actual knowledge, based on the domain of knowledge under consideration. The range of correlations is large, from $r=-0.07$ to 0.68, with a mean of $r=0.45$. However, only the Business/Management domain failed to show a correspondence that did not meet a $P=0.01$ significance level (for a test of $\rho=0$). That is, the remaining 17 domains showed significantly positive correlations between self-estimates of knowledge and actual knowledge. The largest correlations were found for the Science domains (mean $r=0.52$), followed by Civics (mean $r=0.45$), Humanities (mean $r=0.45$) and Business (mean $r=0.16$). Given that each of the self-report measures was only composed of six Likert-type items, the high

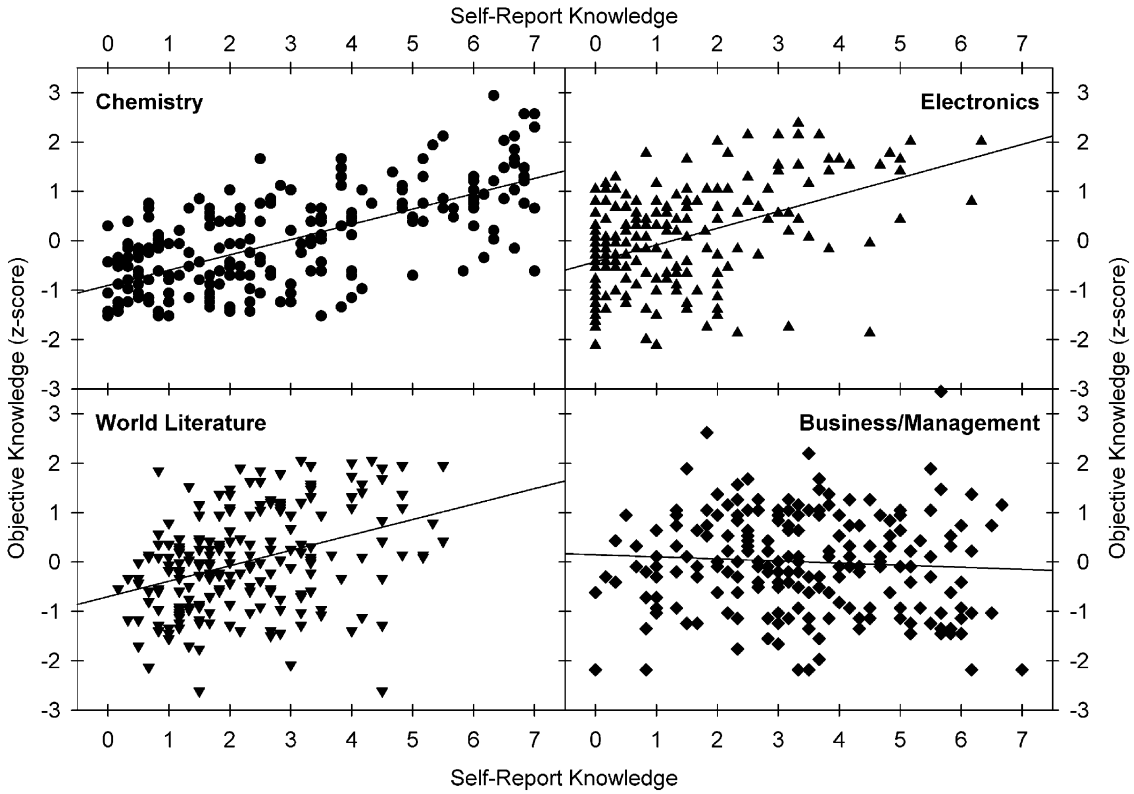

Fig. 3. Scatterplots of self-estimated knowledge plotted against objective knowledge test scores for Chemistry, Electronics, World Literature, and Business/Management. Lines indicate linear regression.

degree of correspondence between the self-report and objective knowledge measures is substantial evidence supporting the notion that participants were reasonably well calibrated in their judgements about their own capabilities, at least in some domains. Fig. 3 provides an illustration of the relations between self-reported knowledge and objective knowledge scores for Chemistry, Electronics, World Literature, and Business.

4.4. Analyses by age, gender, and major field

Fig. 4 presents Cohen's *D*-statistics for the differences in self-reported knowledge and self concept for men and women (negative *D*-statistics correspond to a larger mean difference for men; positive *D*-statistics correspond to a larger mean difference for women). The convention used here is that D-values ranging from 0.20 to 0.49 are considered small effects, 0.5–0.79 are medium and values larger than 0.8 are considered large effects (Cohen, 1988). Women reported significantly higher levels of verbal abilities and clerical abilities, but significantly higher knowledge only in the domain of US Literature. In contrast, men reported significantly higher levels of self-concept for math, spatial, and science abilities, along with 11 of the knowledge scales, most notably in the sciences and civics domains (where the largest effects were found for self-reported electronics and astronomy knowledge). Table 3 shows the correlations between self-report and objective measures by age and gender. As can be seen in the table, no significant gender differences were found for the objective measures of Gf and Gc, but significant advantages were found for men on 10 of the knowledge scales. The only scale with significant advantage for women was Art. Moreover,

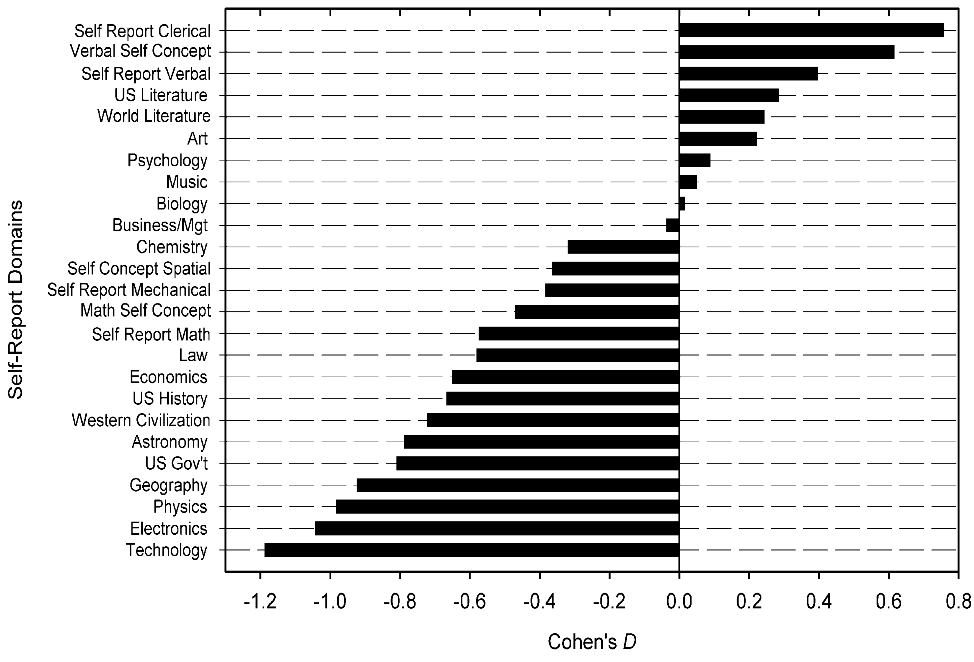

Fig. 4. Cohen's *D*-statistics for the difference between means for men and women on self-reported knowledge and self-concept measures, in S.D. units. Negative *D*-statistics indicate higher average scores for men; positive *D*-statistics indicate higher average scores for women.

there was a close correspondence in the pattern of gender differences in self-reported and objective knowledge. A correlation of $r=0.90$ was found between the 18 pairings of correlations.

Verbal self-concept was positively correlated with age, but math and spatial self-concepts were significantly negatively correlated with age. Similarly, objective measures of Gc showed positive correlations with age, but Gf showed significantly negative correlations with age (in accordance with a large corpus of intellectual ability research). Age correlates in self-reported knowledge were more differentiated than gender differences, with older participants indicating lower knowledge in the Sciences, but significantly higher knowledge in the areas of Civics, Humanities, and Business. The pattern of age effects on objective knowledge scores was largely similar to the self-reported knowledge pattern. A correlation of $r=0.83$ between the respective self-reported vs. objective knowledge correlations was found for the age variable. Thus, patterns of gender differences and age correlates of self-reported and objective scales appear to be quite consistent.

Table 3
Correlations between gender, age for self-report and objective measures

	Gender correlations		Age correlations	
	Self-report	Objective	Self-report	Objective
Verbal self-concept	0.277**		0.086	
Self-estimate of verbal	0.184**		0.247**	
Gc		−0.060		0.143*
Math self-concept	−0.210**		−0.230**	
Self-estimate of math	−0.256**		−0.186**	
Spatial self-concept	−0.169*		−0.167*	
Self-estimate of mechanical	−0.185**		−0.013	
Self-estimate of clerical	0.342**		−0.022	
Gf		−0.107		−0.388**
Astronomy	−0.367**	−0.455**	−0.041	0.014
Biology	0.007	−0.008	−0.260**	−0.202**
Chemistry	−0.150*	−0.225**	−0.273**	−0.240**
Electronics	−0.473**	−0.432**	−0.025	0.200**
Physics	−0.438**	−0.278**	−0.159*	−0.137*
Psychology	0.042	0.103	0.092	−0.099
Technology	−0.508**	−0.352**	−0.056	0.030
US History	−0.301**	−0.408**	0.253**	0.346**
Economics	−0.298**	−0.315**	0.080	0.083
Geography	−0.407**	−0.366**	0.043	0.151*
Western Civilization	−0.326**	−0.326**	0.180**	0.278**
US Government	−0.359**	−0.248**	0.274**	0.219**
Art	0.103	0.173*	0.086	0.224**
Music	0.023	−0.071	0.125	0.297**
US Literature	0.132*	0.016	0.180**	0.357**
World Literature	0.113	0.083	0.137*	0.249**
Business/Management	−0.017	−0.076	0.172**	0.084
Law	−0.268**	−0.090	0.225**	0.189**

Gender coded as "1 = male" and "2 = female"; Gf, fluid intelligence; Gc, crystallized intelligence; $N=228$.
* $P<0.05$.
** $P<0.01$.

Participants were also asked to indicate major field of study (i.e. university major). Of the 228 participants, 208 provided responses that could be classified into one of four categories (Physical Science, Social Science, Humanities, and Business). Although these categories do not map completely onto the major domains of knowledge, they do provide for an illuminating breakdown of both self-report and objective knowledge measures. Fig. 5 shows the breakdown of the self-report and objective knowledge composites by major field of study. Results of a one-way ANOVA for each domain are also shown in the figure. These results show interesting patterns of similarity and difference between the major fields and between knowledge domains. For the Science knowledge domain, participants who majored in the physical sciences reported the highest levels of knowledge, and business majors reported the lowest levels of knowledge. The objective knowledge measures showed the same basic pattern of results for all four major areas. For Civics, no significant differences were found between the major areas in either the self-report or objective measures, though there was a trend for business majors to report higher knowledge levels than other majors, but showed overall (non-significant) lower objective knowledge scores. In the Humanities area, participants who majored in humanities showed the highest self-report knowledge and

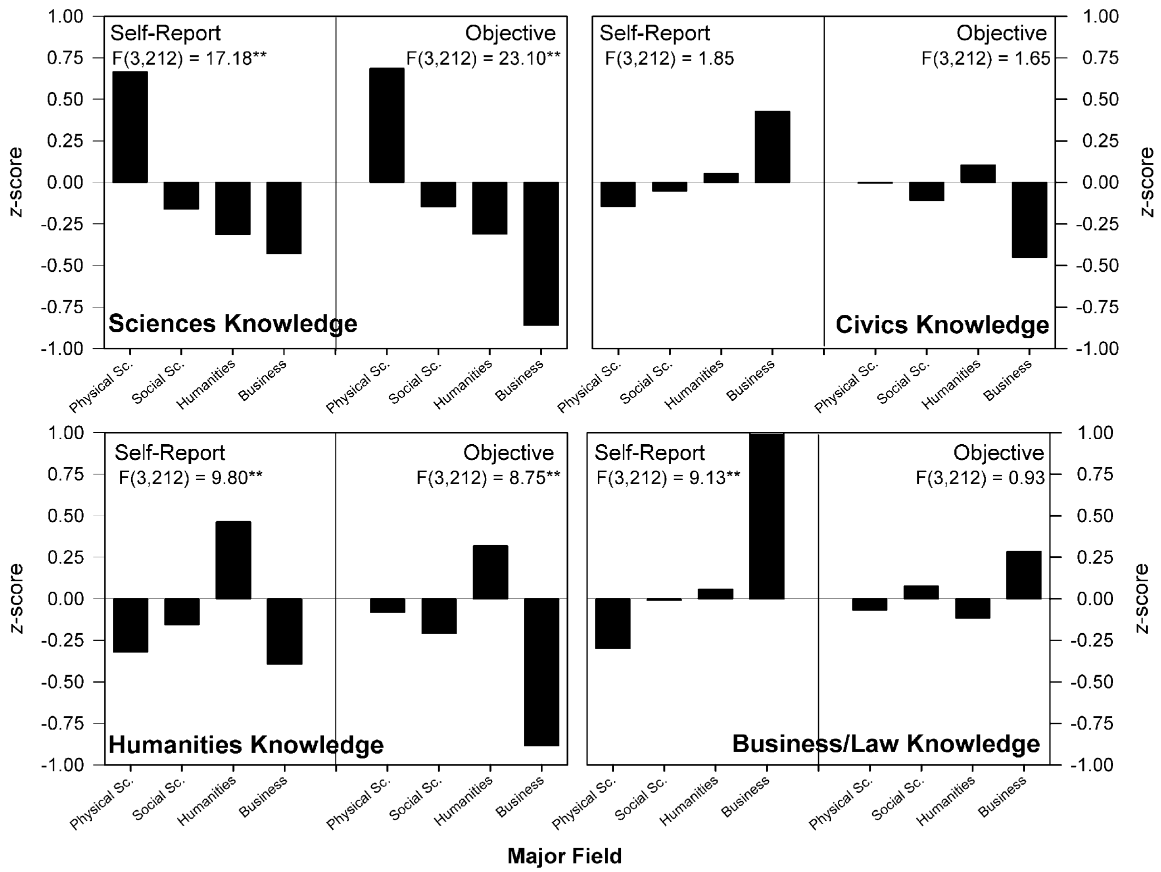

Fig. 5. Mean composite scores (as z-scores) for Self-Report and Objective knowledge measures, by university major field of respondents. F-ratios represent ANOVA performed on scores for Self-Report and for Objective knowledge. $^{**}P<0.01$. Sc, Sciences.

objective knowledge, while business majors showed lowest knowledge in both self-report and objective measures. Finally in the Business domain, business majors reported an extremely high level of knowledge in the self-report section, but showed only a non-significant advantage in the objective knowledge measures. Although a direct comparison across domains is not possible, given lack of measurement invariance, it is interesting to note that the business majors seemed to show the largest discrepancies between self-reported and objective knowledge—*overestimating* their knowledge in all four domains, relative to the other groups.

5. Discussion/conclusions

The results presented here highlight similarities and differences with prior research. For example, like Kruger (1999) and Dunning et al. (1989) we found that broad items and an ambiguous comparison group led to higher self-estimates of ability than when participants rated their abilities using more specific stimuli and an absolute scale. This research also provides additional evidence that individuals tend to polarize their self-concepts of abilities (Marsh, 1990), and that self-assessments for men and women tend to reflect gender stereotypes for abilities (i.e. women reporting higher self-concept in the verbal domains while men reporting higher self-concept in the math and science domains; Marsh & Yeung, 1998).

Most notable, though, when we examined a wide range of abilities and knowledge, we found broad agreement between self-assessed and objective knowledge. We also found evidence for convergent and discriminant validity in self-report measures, indicating that individuals have a general sense of their relative strengths and weaknesses for the abilities and knowledge domains used in this study. Indeed, contrary to other researchers who highlight the tendency for individuals to overestimate their abilities (Kruger & Dunning, 1999; Lichtenstein & Fischhoff, 1977), we found some individuals reporting *no* knowledge on the self-report knowledge scales for 17 of the 18 domains tested.

We also examined and found differences in profiles of self-report and objective knowledge by university majors. For example, Science majors appeared to have a reasonably accurate assessment of their abilities, while Business majors did not. This may be a function of the precision and rigor required for success in science, and perhaps the confidence required for success in business.

The approach and analysis used in this study is less precise than one that asks participants to assess their ability in a narrow domain or their score on a specific test (as has been done in the experimental literature). However, the strength of the approach lies in the breadth of knowledge and abilities studied. The broad approach used in this research allows for more positive conclusions regarding individuals' ability to accurately judge what and how much they know. The correlational approach (see Cronbach, 1957) used also avoided the statistical problems common in research on self-assessment.

Acknowledgements

This research was supported by a grant from the National Institute on Aging, (AG16648), Phillip L. Ackerman, principal investigator. We thank Eric Rolfhus for creating the self-report

knowledge measures, and the Kanfer and Ackerman Laboratory staff for support in data collection and data coding.

References

Ackerman, P. L. (1997). Personality, self-concept, interests, and intelligence: which construct doesn't fit? *Journal of Personality*, *65*(2), 171–204.

Ackerman, P. L. (2000). Domain-specific knowledge as the "Dark Matter" of adult intelligence: Gf/Gc, personality and interest correlates. *Journal of Gerontology*, *55B*(2), P69–P84.

Ackerman, P. L., Kanfer, R., & Goff, M. (1995). Cognitive and noncognitive determinants and consequences of complex skill acquisition. *Journal of Experimental Psychology: Applied*, *1*(4), 270–304.

Ackerman, P. L., & Rolfhus, E. L. (1999). The locus of adult intelligence: knowledge, abilities, and nonability traits. *Psychology & Aging*, *14*(2), 314–330.

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, *49*(6), 1621–1630.

Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, *68*(5), 804–825.

Cantor, N., & Norem, J. K. (1989). Defensive pessimism and stress and coping. *Social Cognition Special Issue: Stress, Coping, and Social Cognition*, *7*(2), 92–112.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684.

DeNisi, A. S., & Shaw, J. B. (1977). Investigation of the uses of self-reports of abilities. *Journal of Applied Psychology*, *62*(5), 641–644.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: the role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, *57*(6), 1082–1090.

Fischhoff, B., Slovic, P., & Lichtenstein, S. (1977). Knowing with certainty: the appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(4), 552–564.

Goff, M. (1994). *Understanding vocational self-concept and related domains: a hierarchical factor-analytic approach*. Unpublished doctoral thesis, University of Minnesota.

James, W. (1890/1950). *The principles of psychology*. NY: Dover Publications.

Kanfer, R., Ackerman, P. L., & Heggestad, E. D. (1996). Motivational skills and self-regulation for learning: a trait perspective. *Learning and Individual Differences*, *8*(3), 185–209.

Kelso, G. I., Holland, J. L., & Gottfredson, G. D. (1977). The relation of self-reported competencies to aptitude test scores. *Journal of Vocational Behavior*, *10*(1), 99–103.

Kruger, J. (1999). Lake Wobegon be gone! The "below-average effect" and the egocentric nature of comparative ability judgements. *Journal of Personality and Social Psychology*, *77*(2), 221–232.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134.

Levine, E. L., Flory, A., & Ash, R. A. (1977). Self-assessment in personnel selection. *Journal of Applied Psychology*, *62*(4), 428–435.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance*, *20*, 159–183.

Lundeberg, M. A., Fox, P. W., Brown, A. C., & Elbedour, S. (2000). Cultural influences on confidence: country and gender. *Journal of Educational Psychology*, *92*(1), 152–159.

Lundeberg, M. A., Fox, P. W., & Puncochar, J. (1994). Highly confident but wrong: gender differences and similarities in confidence judgments. *Journal of Educational Psychology*, *86*(1), 114–121.

Maki, R. H., Jonas, D., & Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin and Review*, *1*(1), 126–129.

Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: theoretical and empirical justification. *Educational Psychology Review*, *2*(2), 77–172.

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: gender differences in development of math and English constructs. *American Educational Research Journal*, *35*, 705–738.

Miner, A. C., & Reder, L. M. (1994). A new look at feeling of knowing: its metacognitive role in regulating question answering. In J. Metcalfe, & A. P. Shimamura (Eds.), *Metacognition: knowing about knowing* (pp. 47–70). Cambridge, MA: MIT Press.

Moreland, R., Miller, J., & Laucka, F. (1981). Academic achievement and self-evaluations of academic performance. *Journal of Educational Psychology*, *73*(3), 335–344.

Rolfhus, E. L., & Ackerman, P. L. (1996). Self-report knowledge: at the crossroads of ability, interest, and personality. *Journal of Educational Psychology*, *88*(1), 174–188.

Rolfhus, E. L., & Ackerman, P. L. (1999). Assessing individual differences in knowledge: knowledge, intelligence, and related traits. *Journal of Educational Psychology*, *91*(3), 511–526.

Shaughnessy, J. J. (1979). Confidence-judgement accuracy as a predictor of test performance. *Journal of Research in Personality*, *13*, 505–514.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: validation of construct interpretations. *Review of Educational Research*, *46*, 407–441.

Wittmann, W. W., & Süß, H. M. (1999). Investigating the paths between working memory, intelligence, knowledge, and complex problem-solving performances via Brunswik Symmetry. In P. L. Ackerman, P. C. Kyllonen, & R. D. Roberts (Eds.), *Learning and individual differences: process, trait, and content determinants* (pp. 77–108). Washington, DC: American Psychological Association.

Yates, J. F., Lee, J. W., & Bush, J. G. (1997). General knowledge overconfidence: cross-national variations, response style, and "reality." *Organizational Behavior and Human Decision Processes*, *70*(2), 87–94.